Center for Mind/ Brain Sciences

Doctoral School in Cognitive and Brain Sciences, 25th cycle

THE NEURAL BASIS OF ACTION UNDERSTANDING

Sandra Petris

Advisor: Dott. Angelika Lingnau

Table of Contents

1 GENERAL INTRODUCTION

In daily life, we constantly observe actions of other people and try to attribute a meaning to these actions. This ability is fundamental to our survivor. Despite intense research, the neural basis of action understanding is still poorly understood. The aim of the present thesis is to investigate the neural networks involved in the process of action understanding in humans using functional magnetic resonance and diffusion tensor imaging (fMRI and DTI).

1.1 Embodied and Disembodied theories of action understanding

One of the major debates in the field of action understanding is related to the role of the sensory-motor system. According to **disembodied accounts**, concepts are abstract mental representations, modality independent, and organized according to conceptual distinctions (e.g., entities versus events). During conceptual processing, sensory and motor information are transformed into abstract and amodal representation of knowledge of past experience. This knowledge would reside in a semantic memory system separate from the brain's modal systems for perception and action (e.g. Mahon & Caramazza, 2008). According to the stronger view of the disembodied account, cognitive functions are organized into independent modules able to process only one specific category of information (e.g. linguistic but not motoric). By contrast, the **embodied account** assumes that knowledge is represented by information that resides within the sensory and motor systems. According to this view, concepts are constituted by the reactivation of sensory-motor information of past experience (Barsalou, 2008; Gallese & Lakoff, 2005). For example, the concept of "running" would activate the motor program for running, the visual image of running and the sound of running and therefore the meaning of running would be distributed among the different sensory-motor systems involved in the experience of running.

Thanks to the discovery of mirror neurons, neurons located in the macaque motor system that are active both during vision and execution of actions (Di Pellegrino, Fadiga, Fogassi, Gallese, &

Rizzolatti, 1992), the debate about the role of sensory-motor systems in conceptual processing has become particularly strong in the domain of action processing. The most popular interpretation of the behavior of mirror neurons, that I will describe later, is that of action understanding and imitation through a **direct matching mechanism** between the visual description of the observed action and the motor representation of that action stored in the observer's motor system (Rizzolatti & Sinigaglia, 2010). According to this view, the observer's motor representation of the observed action would be activated through a **simulation** or "**resonance**" mechanism, in agreement with the embodied cognition account. Whereas the imitation hypothesis, although plausible, has been criticised by the fact that monkeys do not imitate (Hickok & Hauser, 2010), the former has been criticised by several authors claiming that mirror neurons would code the outcome of action understanding rather than being the cause of the process of understanding an action (Csibra, 2008; Hickok & Hauser, 2010; Hickok, 2009; Mahon & Caramazza, 2008). For example Csibra (2008) interprets mirror neurons as a mechanism dedicated to action reconstruction in order to predict and monitor successive motor behavior. This action reconstruction would happen after an inferential process where meaning is assigned.

Mahon and Caramazza (2008) argue that lesion studies may shed light on the role of the motor system in action understanding. In fact, it has been shown that patients impaired in action production are not impaired in action understanding (Negri et al., 2007), suggesting a dissociation between action production and action comprehension, which cannot be reconciled with the hypothesis that mirror neurons subserve action understanding. Before describing the basic findings relative to mirror neurons and the direct-matching account, I will briefly introduce the different levels at which an action can be described.

1.2 Abstract versus concrete representations of actions

One of the major problems in the attempt to study the neural correlates of action understanding is what is meant by action understanding. Actions can be described at lower or higher levels of

representations. A lower level of description includes the kinematic i.e., the trajectory and the velocity profile of the action while higher levels of description include the "what" and "why" of an action, i.e., the goal and the intention. Importantly, the difference between low and high levels of description resides in the fact that low level representations require a one-to-one mapping between the description and the actual action performed while the high level representations refer to a more abstract description, i.e., the semantics of that action (Friston, Mattout, & Kilner, 2011). In other words, an **abstract representation** of an action refers to the level where there is no one-to-one mapping between the description and the actual action performed. These could be the goal and intention levels of an action. For example, the goal of picking up a cup can be achieved with many different actions. On the other hand, a **concrete representation** of an action refers to the unique description of one component of an action, where there is a one-to-one mapping between the description and the actual execution of the action.

1.3 The discovery of mirror neurons in macaques

Mirror neurons were originally discovered in the F5 region, which is part of the inferior frontal cortex of the macaque brain (Di Pellegrino et al., 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996a; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). In F5, two classes of neurons were reported: canonical neurons and mirror neurons. Both mirror and canonical neurons fire when the monkey executes goal-directed actions, such as grasping, holding, tearing and manipulating. Some of these neurons are activated for precision grip, for example, when the monkey grasps a small object with the thumb and index fingers, while others are activated when the monkey grasps big objects, with a whole hand prehension. These neurons, canonical and mirror, also have visual properties. Canonical neurons that fire when the monkey grasps a small object with a precision grip are also activated by the sight of small objects graspable with precision grip but not to the sight of bigger objects graspable with a whole-hand grip. As pointed out by Iacoboni, (2002), *"canonical neurons seem to be coding the affordance of an object, the pragmatic aspect of how-to-grab-that-thing, rather than*

its semantic content". Mirror neurons were reported not to be activated at the sight of the object to be grasped, but, in contrast, to the whole action of grasping performed by the experimenter or another monkey. For example, a mirror neuron in the monkey F5 is activated not only when the monkey grasps a peanut, but also when the monkey observes the experimenter grasping a peanut. The same neuron will not fire at the sight of the object alone and will not fire at the sight of a pantomime of a grasp in absence of the object. Among mirror neurons, two different types were reported, strictly and broadly congruent mirror neurons. The **strictly congruent mirror neurons** fire when the observed and executed action coincide both in terms of goal (grasping) and in terms of how the goal is achieved (grasping with precision grip). On the other hand, **broadly congruent mirror neurons** fire when the goal of the action is the same (grasping) but it is achieved with different means (precision or whole-hand grasp). This latter class of neurons is particularly important since, according to Rizzolatti and Sinigaglia, (2010), they process the information about the goal of an action at different levels of generality.

Regarding the **action goal processing** in monkey F5, Umiltà et al. (2008) trained monkeys to grasp objects with a particular type of pliers. The pliers were built in such a way that grasping could be achieved with opposite sets of movements, i.e., opening or closing the hand. The authors were interested in understanding whether mirror neurons fire in response to the goal of the action (grasping) irrespective of the movement of the hand (opening or closing of the hand). What the authors found was that indeed mirror neurons were active irrespective of the movements, suggesting therefore that mirror neurons code for the goal of the action.

Mirror neurons have also been found in the parietal area of the monkey brain that corresponds to the inferior parietal lobule (IPL) in human brain. Fogassi et al. (2005) recorded cells activity from the IPL neurons in 2 different conditions: in one condition the monkey had to grasp a piece of food and bring it to the mouth, in the other the monkey had to grasp an object and place it in a container. The monkey had to execute and observed the same actions performed by the experimenter. The authors

found that the recorded mirror neurons were differentially activated with respect to the motor act subsequent of the grasping action. For example, some neurons discharged in response to grasping-for-eating but not for grasping-for-placing and others neurons showed the reverse behavior. The same result was found even when the object to be grasped was identical in the 2 conditions and even when the trajectory of the movement was very similar between the 2 conditions. The authors interpreted the behavior of the neurons as coding the intention of the individual performing the action.

To complete this picture about action recognition in monkeys, neurons in the superior temporal sulcus (STS), although lacking mirror and motor properties, respond to the sight of moving biological stimuli, such as hands, faces, bodies (Perrett et al., 1989; Perrett, Mistlin, Harris, & Chitty, 1990). These neurons seem to respond to moving bodies and body parts only when the body or body part is engaged in goal-oriented actions. STS is not part of the mirror network, however it has been suggested to provide higher order visual input of action representation to the mirror network.

In summary, mirror neurons have been discovered in premotor and parietal areas of the monkey brain while STS would provide higher order visual information of action representation to the mirror network.

1.4 Evidence in favor of the direct-matching account in humans

Many attempts have been done to show evidence of brain areas with similar properties as the mirror neurons in monkeys. Using fMRI adaptation, Lingnau, Gesierich, & Caramazza, (2009) wanted to test the hypotheses that cortical areas that contain mirror neurons should adapt if the same motor act is repeated, irrespective of whether this motor act is repeated within the same (‘‘within-modality adaptation’’) or across different modalities (‘‘cross-modal adaptation’’). The adaptation paradigm exploits the physiological property of a neuronal population to adapt, i.e., to decrease the firing rate in response to a repeated exposure to that stimulus. Therefore mirror neurons should show a cross-

modal adaptation in response to observation and execution of actions. To test this hypothesis, participants were scanned in two conditions: an observed condition where participants had to observe movies of a hand performing meaningless actions and an execution condition where participants had to execute with their hand the actions. The authors found adaptation within each condition, observation and execution, but critically they did not find cross-modal adaptation when the action was first executed and then observed. This finding led the authors to challenge the claim that the direct-matching account exists in humans.

Several other studies used fMRI adaptation to study cross-modal adaptation, producing mixed results. Chong, Cunnington, Williams, Kanwisher, and Mattingley (2008) presented participants with videos of pantomimed hand actions such as hammering a nail or shot gun. Participants had to either observe and execute or execute and observed the actions. The authors found adaptation for the execution-observation condition in right IPL leading to the conclusion that this region contains neurons with mirror properties. Dinstein, Hasson, Rubin, and Heeger (2007) had participants playing the rock-paper-scissor game against a videotaped opponent. Observed movements (performed by a virtual opponent) and executed movements (performed by the subject) were sometimes repeated on consecutive trials. The authors compared the responses to trials in which a movement was repeated (e.g., rock preceded by rock) with trials in which it was not (e.g., rock preceded by paper). They found numerous cortical regions showing adaptation to repeats, either observed or executed actions, and more importantly several cortical areas showed an overlap between observed and executed actions such as anterior inferior frontal sulcus (aIFS), ventral premotor (vPM), anterior intraparietal sulcus (aIPS), superior intraparietal sulcus (sIPS), posterior intraparietal sulcus (pIPS), and lateral occipital (LO). However, they found no signs of crossmodal adaptation.

Kilner, Neal, Weiskopf, Friston, & Frith (2009) had participants either observing or executing hand goal-directed actions. These actions consisted in grasping a nail with index and thumb and pulling a

ring with the index. An fMRI-compatible manipulandum that allowed the subject to make those grip actions was built and place inside the scanner. For the observation condition subjects had to observed video clips of a right hand performing the very same grip actions. The authors were interested in finding whether human IFG showed cross modal adaptation adopting goal-directed actions. They reasoned that the failing of finding cross-modal adaptation in several previous studies was due to the lack of goal-directed actions, since goal directed actions were the only actions able to trigger mirror neurons activation in monkey F5. The authors indeed found cross-modal adaptation both when an observed action was followed by an executed action and vice versa.

A recent study by Oosterhof, Tipper, & Downing (2012) adopted multi-voxel pattern analyses (MVPA) to test the hypothesis of whether action representations in PMv generalize across the perspective (first- and third person), irrespective of the modality (action observation, action execution). The authors reasoned that in order to understand other people's actions, they need to observe the action performed in a third person perspective while the majority of the studies in the literature employed a first person perspective during action observation. In addition it is worth noting that mirror neurons in monkey brain respond to the experimenter's actions who is in front of the animal assuming therefore a third person perspective. Oosterhof et al. (2012) scanned participants during both observation and execution of goal-directed actions. They build an fMRI compatible table. On the top of it, a cup-shaped object was attached with an elastic string. The table was positioned over the navel of the participants. The task consisted in either observing or executing one out of two possible actions (lift/ slap the cup). During observation, the first person condition consisted in videos of a model in supine position with the table and object placed similarly to the participants in the scanner. By contrast, the third-person perspective condition was constituted by videos of the same model with a similar position of the table and object, but with the model standing on the opposite side of the table. The authors showed that actions perceived from a first person perspective showed cross-modal coding in PMv, but the same actions, perceived from a third person perspective did not. By contrast, more posterior parietal and occipitotemporal regions

showed perspective invariance. The data by Oosterhof et al. (2012) thus suggest that posterior parietal and occipitotemporal cortex, contain view-invariant cross-modal action representations, i.e. properties typically assigned to mirror neuron areas.

The studies reviewed so far aimed to investigate whether signs of mirror neurons existence were present in humans using indirect measures such as fMRI. However, an attempt to demonstrate directly the existence of mirror neurons in humans was done by Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried (2010). They recorded extracellular activity from patients with pharmacologically intractable epilepsy. Patients were implanted with intra-cranial depth electrodes to identify seizure foci for potential surgical treatment. Patients were required to either observe or execute actions. In the observation condition they observed video clips depicting an hand grasping a mug with either precision grip or whole-hand prehension. In the execution condition, they had to perform the same observed actions after a cue appearing on the screen. The authors found cells responding to both the observation and execution conditions in medial frontal and medial temporal lobe, namely, hippocampus, parahippocampal gyrus, and entorhinal cortex. It is important to note that these areas are not part of mirror neuron system neither in monkeys nor in humans and therefore it is difficult to reconcile these data with data from the macaque literature.

Although the above data do not provide convincing evidence for the existence of cortical areas in humans with properties similar of those of mirror neurons reported in macaques, a series of TMS and fMRI studies have been performed to demonstrate either the resonance mechanism or the involvement of the putative human mirror system in coding the goal of actions.

The first attempt to show evidence that a resonance mechanism was at work during action observation was done in a TMS study by Fadiga et al. (1995). In this experiment, TMS pulses were delivered over the motor cortex of healthy participants while observing the experimenter performing goal-directed actions (grasping an object) and arm movements (tracing in the air complex geometrical shapes), and as a control condition participants had to observed objects. Motor

evoked potentials (MEP) were recorded from hand muscles. The authors found increased MEPs during movement observation relative to object observation. Moreover, the MEP pattern reflected the pattern of muscle activity recorded when the subjects executed the observed actions. This study has been taken as support for the direct-matching mechanism in humans. Since MEPs are the result of stimulation of motor neurons in primary motor cortex, it is interesting to note that the resonance mechanism could be at work also in primary motor cortex which was not considered as part of the mirror network in monkeys.

In order to test whether the human mirror neuron system coded for the goal of an action, Cattaneo, Caruana, Jezzini, and Rizzolatti (2009), in analogy with the monkeys experiment, applied TMS pulses over the motor cortex of volunteers while they were watching the experimenter performing grasping actions with two different pliers: the "normal pliers" were opened by extension of the fingers and closed by their flexion while the "reverse pliers" used the opposite pattern of movements. In one condition participants watched the experimenter opening and closing the pliers while in another condition participants watched the experimenter performing goal-directed actions like grasping an object with the two different pliers. In an additional condition, subjects had to imagine themselves operating with the pliers. Motor evoked potentials were recorded from the opponent pollicis (OP) muscles of the hand of participants. During observation of actions devoid of a goal, the MEP amplitudes, regardless of pliers used, reflected the muscular pattern involved in the execution of the observed action. By contrast, during the observation of goal-directed actions, the MEPs from OP were modulated by the action goal, increasing during goal achievement despite the opposite hand movements necessary to obtain it. During motor imagery, the MEPs recorded from OP reflected the muscular pattern required to perform the imagined action. The authors concluded that covert activity in the human motor cortex may reflect different aspects of motor behavior. Imagining oneself performing tool actions or observing tool actions devoid of a goal activates the representation of the hand movements that correspond to the observed ones. By contrast, the observation of tool actions with a goal incorporates the distal part of the tool in the observer's body

schema, resulting in a higher-order representation of the meaning of the motor act.

In another study, Cattaneo, Sandrini, & Schwarzbach (2010) used TMS adaptation (TMSa) to investigate whether frontal and parietal areas contain abstract representations of observed actions. In TMSa, prior to TMS stimulation, cortical areas specific for a given cognitive stimulus are perceptually adapted by repeated presentation of the same stimulus. This procedure induces habituation in a subset of cells, making them a selective target for TMS. Once TMS is applied over the habituated area, it is assumed to selectively improve performance in processing the adapted stimulus. In the experiment by Cattaneo et al. (2010), participants were presented with adapting movies followed by the presentation of a series of static test pictures. Participants were required to carefully watch the adapting movies that repeatedly showed a single motor act (pushing or pulling) performed by a hand or a foot. They were then required to respond as fast and as accurately as possible to static test pictures whether the depicted motor act was same or different from the one seen in the adaptation movie. TMS pulses were applied at the onset of each test picture and were applied over the left and right PMv, left IPL and left STS. Reaction times were recorded. The authors found that TMS over the left and right PMv and over the left IPL induced a selective shortening of reaction times to stimuli showing a repeated (adapted) action, regardless of the effector employed while TMS applied over the left STS induced shortening of RTs for adapted actions but only if also the effector was repeated. The authors concluded that a hierarchy in the action processing exists with a degree of abstraction in processing actions from STS coding low level features of the action to parieto-frontal network coding for abstract representation of actions. Following the same line of evidence, Gazzola, Rizzolatti, Wicker, and Keysers (2007) measured brain activity while volunteers were instructed to observe video clips in which either a human or a robot arm performed grasping actions with different objects (e.g. grasping a glass of wine or grasping a spoon, etc.). The rationale of this study was to investigate whether an abstract classification of an action, such as the ‘what’ of an action is coded by the hMNS irrespective of the ‘how’ of an action (e.g. the agent or the specific kinematic of an action). In other words, if the

activity in the hMNS is similar for the different agents, either the human or robotic arm, the hMNS would primarily match the goal of the action. The authors found similar brain activity for either the robot or human arm. They concluded that the hMNS codes abstract features of an action such as the goal.

Taken together, all these data provide evidence that the motor system is involved not only in action observation but also in coding higher-order representation of action such as the goal of an action.

1.5 Empirical evidence challenging the direct matching account

The direct matching account of action understanding has been recently challenged on the basis of the argument that activation in the motor system during action observation might be triggered by the outcome of action understanding, and that cognitive processing that underlies action understanding takes place in structures outside the motor system (Csibra, 2008; Hickok, 2009; Lingnau et al., 2009; Mahon & Caramazza, 2008).

Several studies have tested the hypothesis of the existence of an amodal center responsible for the processing of concepts independent of the sensory-motor properties that a concept may contain. For example, Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, (2008) presented participants with pairs of either nouns or verbs belonging to different categories such as actions (i.e., kicking) or non-actions (i.e., thinking) and asked participants to rate how related in meaning the words were. Pairs of words were presented in the hearing modality and brain responses were collected through fMRI. The authors were interested in finding areas of the brain active irrespective of the perceptual-motor properties a word may refer to. According to the disembodied theory of cognition this brain area should be equally activated for both kicking and thinking. The authors found areas of the brain located in posterior lateral temporal cortex (PLTC), an area previously associated with the coding of action verbs and therefore suggested to be involved in the visual representation of concepts in virtue of the fact that PLTC is located near higher order visuo-motion areas such as middle temporal (MT) and superior temporal sulcus (STS) that code for visual aspects of motion (Martin, Wiggs,

Ungerleider, and Haxby, 1996). However, Bedny et al., (2008) found that PLTC responds equally to actions and non-action categories of words leading to the suggestion that PLTC may represent an amodal brain region that reflect the retrieval of event concepts. However, PLTC was found also to have a preference for the grammatical class of the words presented, i.e., verbs vs nouns irrespective of the belonging category, so it is not clear whether this region codes for concepts or grammatical class. Whereas Bedny et al. (2008) could not disentangle between grammatical class and concepts processing, Bedny, Caramazza, Pascual-Leone, and Saxe (2011) tested the hypothesis of whether left middle temporal gyrus (MTG) is a brain region for the storage of visual representations related to actions verbs or whether it stores abstract representations of action verbs irrespective of their sensory components. In order to test these hypotheses, the authors measured the BOLD signal in congenitally blind and sighted adults while they performed a semantic judgment task with action verbs as well as 2 other categories of verbs (mental and change of state verbs) and 3 categories of nouns that varied in visual-motion features (animals, artifacts, natural inanimate objects). If action concepts are partly visual, then the concepts of congenitally blind individuals should be altered in that they lack these visual features. The authors found that the functional profile and location of MTG was identical in sighted and congenitally blind individuals, so they concluded that the left MTG contains abstract representations of verb meanings rather than sensory representations related to action verbs developed through experience.

Other studies suggest the MTG, which is part of PLTC is a central node for conceptual processing. Voxel-based lesion symptom-mapping (VBLSM) is a technique to relate behavioural function to anatomy on a voxel by voxel basis in patients with a brain lesion (Bates et al., 2003; Rorden & Karnath, 2004). Using this technique, Kalenine, Buxbaum, & Coslett (2010) reported data from 43 left hemisphere stroke patients in a semantic and spatial action recognition tasks. In the semantic recognition task, patients heard and saw an action verb (e.g. ‘hammering’) and then had to select from two video clips of an experimenter performing either “hammering” or another action which had a semantic relation to the first action, such as “sawing” (both hammering and sawing are tool-

related actions), the one corresponding to the verb. The spatial recognition task was identical to the semantic one except that the difference between the pairs of video clips showing the experimenter performing the actions were that an error in the hand posture, arm posture or amplitude/timing components was present. The authors reported that performance in the semantic and spatial recognition tasks was associated with lesioned voxels in the posterior MTG and IPL, respectively. They concluded that the posterior MTG appears to serve as a central node in the association of actions and meanings. In other words, MTG might be seen as a central hub in the network involved in action semantics where action concepts *"abstract away from the specific sensory and motor events that have gone along with the instantiation of those concepts in the past"* (Mahon & Caramazza, 2008).

In the following two studies reported in this thesis I investigated whether areas outside the mirror network are modulated by understanding of the goal of an action in comparison to identifying the effector that constitutes an action (Study I). One problem in the direct-matching account of action understanding is that the available data can be accommodated by the disembodied view of cognition, in that the involvement of the motor system in action processing could be the outcome rather than the processing of action meaning, with the action meaning being processed outside the motor system. By contrast, it would be problematic to explain the recruitment of areas outside the mirror network in tasks that require action understanding on the basis of embodied accounts of action understanding. The first experiment aims to test the hypothesis of whether areas inside or outside the motor system are recruited during tasks that require action understanding. The second study aims to further investigate the areas involved in action understanding by performing a diffusion tensor imaging (DTI) study in order to localize white matter bundles involved in action understanding through correlation between behavioral performance in tasks used in the first study and structural information of the brain.

2. STUDY I: ACTION UNDERSTANDING WITHIN AND OUTSIDE THE MOTOR SYSTEM: THE ROLE OF TASK DIFFICULTY

Abstract

When we observe actions, we activate parietal and premotor areas that are also recruited when we perform actions ourselves. It has been suggested that this action mirroring is causally involved in the process of action understanding. Alternatively, it might reflect the outcome of action understanding, with the underlying cognitive processes taking place elsewhere.

To identify and characterize areas involved in action understanding, we presented participants with point-light displays depicting human actions and engaged them in tasks that required understanding the effector (Effector Task) or the goal (Goal Task) of an action. We observed a stronger blood-oxygen level dependent signal during the Goal in comparison to the Effector Task not only in parietal and premotor areas, but also in the middle temporal gyrus (MTG) and the anterior ventrolateral prefrontal cortex. In MTG, the Goal Task led to a higher signal than the Effector Task only when actions were easy to understand, whereas frontal areas showed this difference also when the task was difficult, a finding that is not caused by a ceiling effect. Our results suggest an interplay between temporal and frontal areas that is modulated by task difficulty and thus provide important constraints for biologically plausible models of action understanding.

2.1 Introduction

Successful interactions with our environment depend on our ability to interpret and understand actions performed by others. Motor theories of action understanding suggest that this ability crucially depends on the recruitment of the same structures that are involved when we perform the same actions ourselves. According to this view, we understand actions by means of a direct matching mechanism between observed actions and actions that are stored in the observer's motor system (Rizzolatti & Craighero, 2004; Giacomo Rizzolatti & Sinigaglia, 2010). An alternative account that we will refer to as non-motor theories of action understanding assumes that activation in the motor system during action observation might be triggered by the outcome of action understanding, and that cognitive processing that underlies action understanding takes place in structures outside the motor system (Csibra, 2008; Hickok, 2009; Lingnau et al., 2009; Bradford Z Mahon & Caramazza, 2008).

Neurophysiological studies have provided evidence that is compatible with motor theories of action understanding. In particular, a subset of neurons in macaque inferior premotor cortex has been reported to be activated both during the observation and execution of specific actions (Di Pellegrino et al., 1992; V Gallese et al., 1996a). Some of these neurons are sensitive to the type of action (e.g. grasping an object), irrespective of the way these actions are performed (e.g. using normal or reversed pliers; Umiltà et al., 2008). Moreover, these neurons respond during hand-object manipulations even if the final part of the action is hidden from the monkey and thus can only be inferred (Umiltà et al., 2001). In humans, the inferior parietal lobe (IPL) and the inferior frontal gyrus (IFG) have been suggested to form the putative human mirror system (Buccino et al., 2001; Giacomo Rizzolatti & Craighero, 2004), but it is debated if these areas contain neuronal populations with properties similar to those reported in macaques (Chong TT et al., 2008; Dinstein I, JL Gardner et al., 2008; Dinstein I et al., 2007; Dinstein I, C Thomas et al., 2008; Hickok G, 2009; Kilner JM et al., 2009; Lingnau A et al., 2009; Mukamel R et al., 2010; Oosterhof NN et al., 2011).

Whereas both motor and non-motor theories of action understanding are compatible with an involvement of premotor regions in task requiring action understanding, only motor theories of action understanding predict that an impairment of the motor system affects the ability to understand actions. Studies testing this prediction have produced mixed results. In line with motor theories of action understanding, several studies report that performance in tasks requiring gesture production and recognition are correlated in groups of patients suffering from a lesion in the left hemisphere (Buxbaum, Kyle, & Menon, 2005; Pazzaglia, Smania, Corato, & Aglioti, 2008). However, whereas such correlations can be found at the group level, there are important exceptions from these patterns. As an example, Negri et al. (2007) described several patients that were impaired for object use but normal for action recognition. Moreover, they reported one patient (P.T.) with impaired imitation of both pantomimes and intransitive actions, but normal performance in pantomime recognition and object use (see also Mahon & Caramazza, 2005□). Likewise, Papeo, Negri, Zadini, & Rumiati (2010) demonstrated double dissociations between action naming and imitation.

Motor and non-motor theories of action understanding differ in one additional aspect. Whereas motor theories assume that the motor system contains high-level representations of action goals that need to be recruited to understand actions, non-motor theories assume that such high-level representations can be found outside the motor system. This comparison is thus an example of 'Forward Inference' (Henson, 2006) since non-motor, but not motor theories of action understanding, predict the involvement of areas outside the motor system.

Studies testing this prediction have produced conflicting results. Hamilton & Grafton, (2006) reported that aIPS is sensitive to the type of object being grasped, but not to the spatial location of the object, and conclude that aIPS represents the goal of an action. Using state-dependent TMS, Cattaneo et al. (2010) reported that PMv and IPL is sensitive to the type of action (push, pull) irrespective of the effector (hand, foot), whereas STS was sensitive to both manipulations,

suggesting a more abstract representation in PMv and IPL in comparison to STS. By contrast, neurons in STS show a generalization across a variety of manipulations such as viewing angle and movement velocity (Perrett et al., 1989, 1990), indicating that they might represent the goal of an action. In line with this view, Lestou, Pollick, & Kourtzi, (2008) reported that STS is sensitive to differences in the type of action, irrespective of the kinematics, whereas PMv is sensitive both to movement kinematics and the type of action. Tranel, Kemmerer, Adolphs, Damasio, & Damasio (2003) reported that retrieval of action knowledge is impaired not only by lesions in the left premotor/ prefrontal cortex and parietal cortex, but also by lesions in the left posterior middle temporal region.

A problem that is common to many previous studies is that the tasks often can be performed without having to understand the actions, e.g. on the basis of judgments of the spatial configuration of stimuli (Urgesi, Candidi, Ionta, & Aglioti, 2007; Urgesi, Moro, Candidi, & Aglioti, 2006) or on the basis of the distinction between correct and incorrect actions (Pazzaglia et al., 2008). Such differences might simply be due to familiarity or salience and thus do not necessarily reflect the process of action understanding (see also Kalenine et al., 2010).

In the current study we aimed to identify and characterize the network of areas that are involved if participants have to understand the goal of an action in comparison to identifying the effector that constitutes an action while keeping visual stimulation identical, thereby preventing confounding factors such as familiarity or salience. According to motor theories of action understanding, we should find that understanding the goal versus the effector of an action reveals areas within parietal and premotor cortex, in particular the IPL and the PMv. By contrast, if activation within parietal and premotor regions during action understanding reflects the outcome rather than the process underlying action understanding, we expect to find differences between the goal and the effector task also outside the motor system. To test our predictions, we presented participants with point-light displays depicting human actions (Throwing a Ball, Punching Someone, Kicking a Ball,

Kicking Someone) while engaging them in three tasks requiring different levels of action understanding (Figure 1A). During the Red Dot Task, participants had to indicate whether or not one of the points forming the actions had briefly turned red. During the Effector Task, participants had to indicate if the relevant effector was the arm (as in throwing a ball) or the leg (as in kicking a ball). During the Goal Task, participants had to indicate whether the point-light display depicted an action involving a ball (as in throwing a ball) or not (as in punching someone). To be able to examine if areas revealed by these tasks are sensitive to task difficulty, we parametrically varied this factor by applying different levels of spatial noise to the point-light displays (Figure 1B).

Our paradigm revealed a network of areas that showed a stronger blood-oxygen level dependent (BOLD) response during the Goal Task in comparison to the Effector Task. These areas included, but were not restricted to parietal and premotor areas. In particular, our results suggest an important role of the middle temporal gyrus (MTG) and the anterior ventrolateral prefrontal cortex (aVLPFC) in action understanding. These results are compatible with non-motor, but not with motor theories of action understanding and thus have important implications for the assertion that higher level cognitive processes are mediated by a direct, not cognitively mediated simulation within the motor system.

2.2 Materials and Methods

Participants. 17 participants (9 males, 7 females, mean age 28.9 years) participated in this experiment. Vision was normal or corrected-to-normal using MR-compatible glasses. All participants were neurologically intact and gave written informed consent for their participation. The experimental procedures were approved by the ethical committee for research involving human subjects at the University of Trento.

Procedure and visual stimulation. Stimuli consisted of point-light displays recorded from two different actors (one male, one female) using a motion-capture device (Qualisys) at a sampling

frequency of 250 Hz outside the MR lab. We used 13 reflective markers attached to the head, shoulders, elbows, hands, hips, knees, and feet. Four different types of human actions (throwing/ kicking a ball; punching/ kicking someone; Figure 1A) were used. Ten different exemplars of each action were produced by each actor. To make sure that actors performed the different exemplars of each action with similar timing, and to make sure that the actions were performed in a comparable way across the two actors, we provided the actors with a prototype video of each action via a laptop in front of them during stimulus generation. Coordinates of each point-light stimulus were reconstructed using software available from Qualisys. For each stimulus, five different viewing angles (lateral to the left/ right, half-frontal to the left/ right, frontal) were reconstructed. Each point-light display was downsampled to 60Hz and cut to a length of 1.5 s using software written in MATLAB.

Difficulty of action understanding was manipulated by either using the original trajectories (low noise level), or by rotating the trajectories of 6 (medium noise level) or 12 (high noise level) numbers of markers by either 90, 180 or 270° (Saygin, Wilson, Hagler Jr., Bates, & Sereno, 2004)□. This rotation disrupts local form information while keeping the overall physical stimulation intact. The advantage of this procedure, in contrast to paradigms used in previous studies, is that it allows quantification of performance in action understanding.

In each trial, participants were presented with a 1.5 s point-light display depicting one of the actions. At the same time, they were asked to perform tasks that require different levels of action understanding (Figure 1A). During the Red Dot Task, participants were asked to indicate whether or not one of the 13 markers briefly turned red. During the Effector Task, they were asked to judge whether the action mainly involved the arm (throwing a ball/ punching someone) or the leg (kicking a ball/ kicking someone). The Goal Task required to indicate if the action involved a ball (throwing/ kicking a ball) or not (punching/ kicking someone).

Stimuli were back-projected onto a screen by a liquid crystal projector at a frame rate of 60 Hz and

a screen resolution of 1024 x 768 pixels (mean luminance, 109 d/m2). Participants viewed the stimuli binocularly through a mirror above the head coil. The screen was visible as a rectangular aperture of 17.8 x 13.4°.

The size of each point-light display was normalized to appear on the screen with a width and height of 5.25°. The size of each single marker was 0.14°.

In half of all trials, one of the markers (randomly selected) turned red for four consecutive frames (66.7 ms), whereas in the other half of the trials, all markers were white throughout the entire trial. The duration of the red dot was chosen on the basis of behavioral pilot data to make sure that the Red Dot Task was demanding. To prevent missing the red marker when it was presented very close in time to the onset or the offset of the point-light display, it always appeared within a time window of 500 to 1000 ms after the onset of the trial.

Stimulus presentation, response collection and synchronization with the scanner were controlled with in-house software (Schwarzbach, 2011), based on the MATLAB Psychtoolbox-3 for Windows (Brainard, 1997).

2.3 Instructions and training

Before the experiment, participants were given written instructions explaining their task, followed by several practice blocks until the participant was able to perform the tasks properly. For the Red Dot Task, participants had to indicate by button press if one of the markers had briefly turned red (left: yes, right: no). During the Effector Task, participants had to indicate if the action mainly involved the arm (left button) or the leg (right button). During the Goal Task, they were instructed to indicate if the action involved the use of a ball (left: yes, right: no).

2.4 Design

The experimental design is illustrated in Figure 1C. We used a mixed design, with task blocked (15 s, followed by 12 s rest). Within each block, noise level was varied from trial to trial, with each

noise level occurring two times per block, leading to six trials per block, for a total of 24 repetitions of each noise level per scanning run. The type of action was assigned randomly, with each type of action occurring 24 times within each scanning run.

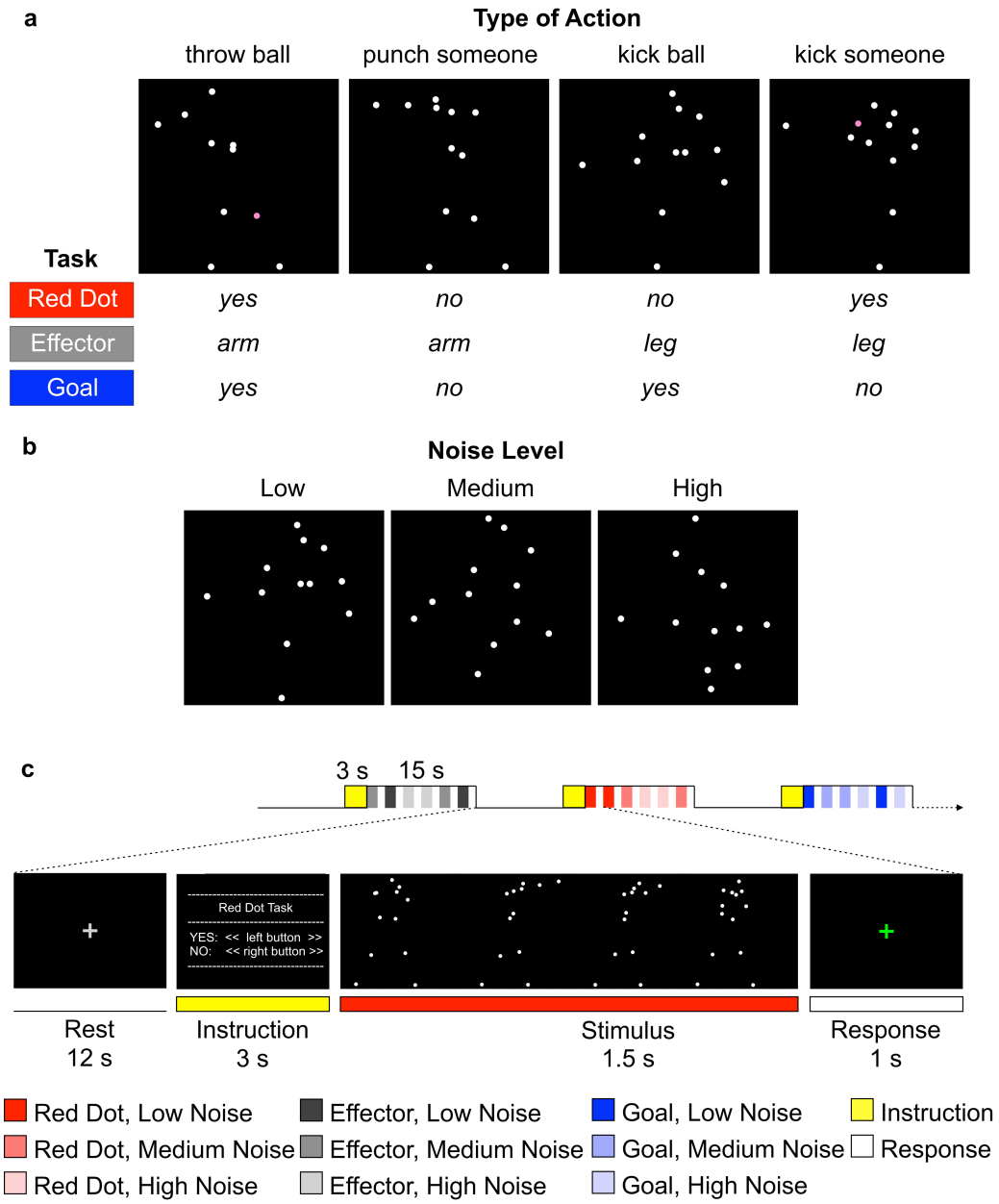


Figure 1. *A: Stimuli consisted of point-light displays depicting four different actions (throw ball, punch someone, kick ball, kick someone). During the Red Dot Task, participants had to indicate whether or not one of the 13 markers had briefly turned red. During the Effector Task, participants had to indicate if the relevant effector was the arm (as in throwing a ball) or the leg (as in kicking a ball). During the Goal Task, participants had to indicate whether the point-light display depicted an action involving a ball (as in throwing a ball) or not (as in punching someone). B: Difficulty of action understanding was parametrically manipulated by applying different degrees of noise to the point-light displays (Saygin et al., 2004). C: Experimental design. The experiment consisted of 8 scanning runs of 6.25 min, with each run consisting of 72 trials. We used a mixed design, with task blocked (15 s, followed by 15 s rest). Within each block, noise level was varied from trial to trial, with each noise level occurring two times per block, leading to six trials per block, for a total of 24 repetitions of each noise level per scanning run. The type of action was assigned randomly, with each type of action occurring 24 times within each scanning run.*

Each task block started with a brief (3s) written instruction of the upcoming task, followed by six trials, lasting 1.5 seconds each. Each single trial was followed by a 1 sec ITI during which a green fixation cross was shown. Participants were instructed to respond while the fixation cross was present.

Within each scanning run, no point-light display was shown more than once, with type of actor and viewing angle assigned randomly. The order of blocks followed a regular sequence (e.g. ABC), with each participant having two different sequences in odd and even runs (ABC versus BCA for participants 1, 2, 7, 8, 13 and 14; CAB versus ACB in participants 3, 4, 9, 10, 15, 16; BAC versus CBA in participants 5, 6, 11, 12, and 17).

In total, we created 4 (action type) x 2 (actor) x 10 (exemplars of each action type per actor) x 5 (viewing angle) x 3 (noise level) = 1200 different point light displays.

Altogether, there were 36 conditions (4 types of actions x 3 tasks x 3 noise levels). However, for data analysis, we aimed to average across the type of action, thus leaving 9 conditions (3 tasks x 3 noise levels), with 64 repetitions per condition per participant. To identify regions of interest, we computed a contrast comparing the Effector and the Goal Task versus the Red Dot Task (see section "Definition of Regions of Interest"). For statistical analysis of the effect of task and noise level, we therefore left out the Red Dot Task, using 6 conditions (2 tasks x 3 noise levels; see section "Statistical Analysis").

2.5 Data acquisition

We acquired fMRI data using a 4 T Bruker MedSpec Biospin MR scanner and an eight-channel birdcage head coil. Functional images were acquired with a T2*-weighted gradient-recalled echo-planar imaging (EPI) sequence with fat suppression (FS). Before each functional scan, we performed an additional scan to measure the point-spread function (PSF) of the acquired sequence, which serves for correction of the distortion expected with high-field imaging (Zaitsev, Hennig, &

Speck, 2004).

We used 31 slices, acquired in ascending interleaved order, slightly tilted to run parallel to the calcarine sulcus [time to repeat (TR), 2250 ms; voxel resolution, 3 x 3 x 3 mm; echo time, 33 ms; flip angle (FA), 76°; field of view (FOV), 192 x 192 mm; gap size, 0.45 mm]. Each participant completed 8 scanning runs of 168 volumes each.

To be able to coregister the low-resolution functional images to a high-resolution anatomical scan, we acquired a T1-weighted anatomical scan (magnetization-prepared rapid acquisition gradient echo; voxel resolution, 1 x 1 x 1 mm; FOV, 256 x 224 mm; generalized autocalibrating partially parallel acquisitions with an acceleration factor of 2; TR, 2700 ms; inversion time, 1020 ms; FA, 7°).

2.6 Data analysis

Behavioral performance in the three tasks was analyzed using d' (Macmillan & Creelman, 1991), which corrects for response bias. Data analysis of anatomical and functional data, including cortex segmentation and inflation, was performed using BrainVoyager QX 2.2 (BrainInnovation) in combination with the BVQX Toolbox and custom software written in Matlab (MathWorks).

Preprocessing. To correct for distortions in geometry and intensity in the echo-planar images, we applied distortion correction on the basis of the PSF data acquired before each EPI scan (Zeng & Constable, 2002). Before additional analysis, we removed the first four volumes to avoid T1 saturation. We aligned the first volume of the first run, which was closest in time to the acquisition of the anatomical scan, to the high-resolution anatomy (9 parameters). Next, we performed 3D motion correction with trilinear interpolation using the first volume of the first run of each participant as reference, followed by slice timing correction with ascending interleaved order. Functional data were temporally high-pass filtered using a cutoff frequency of three cycles per run. Furthermore, we applied spatial smoothing with a Gaussian kernel of 8 mm full-width at half-

maximum. For group analysis, both functional and anatomical data were transformed into Talairach space using trilinear interpolation.

Cortex reconstruction. Within each hemisphere, the border between gray and white matter was segmented and reconstructed. The resulting surfaces where then smoothed and inflated.

Definition of regions of interest. To identify regions of interest, we aimed to choose a contrast that allows identifying areas that are involved in action understanding in contrast to passive observation of biological motion, without biasing our data towards a difference between a task that requires a higher level of action understanding (i.e., determining if the action involved a ball or not) in comparison to a lower level task that can be carried out on the basis of lower level aspects of the action (i.e., the type of the effector). To this aim, we computed a random-effects (RFX) general linear model (GLM) including the factors Task (Red Dot, Effector, Goal) and Noise Level (Low, Medium, High). We also included the first derivative of each predictor time course to be able to model shifts of the hemodynamic impulse response function. Furthermore, we included six parameters resulting from from 3D motion correction (x, y, z translation and rotation) in the model. Each predictor time course was convolved with a dual-gamma hemodynamic impulse response function (K. J. Friston et al., 1998). The resulting reference time courses were used to fit the signal time course of each voxel.

To identify ROIs, we computed the contrast [Effector Task, Goal Task] > Red Dot Task, using high noise level trials only. The resulting statistical map was thresholded using a false-discovery rate (FDR) < .01 in combination with a cluster threshold of 4 contiguous voxels.

Statistical analysis. Within each ROI, we computed a RFX GLM analysis with the factors Task and Noise Level and extracted the resulting beta weights separately for each participant. Next, we performed an ANOVA with the factors ROI x Task x Noise Level. Since the comparison between the Red Dot Task and the Effector and Goal Task had been used for ROI identification, we

restricted the ANOVA to the Effector and Goal Task. Additional statistical analyses were carried out if justified by significant interactions.

2.7 Results

Behavioral Data: D Prime. Figure 2A shows the behavioral data (D') collected inside the MR scanner as a function of Noise Level (low, medium, high) and Task (Red Dot, Effector, Goal). Two observations are evident: first of all, as expected, performance decreased with increasing noise level for the Effector Task and the Goal Task, but not the Red Dot Task. Second, performance differed between the three tasks, with highest overall performance for the Effector Task and lowest performance for the Red Dot Task. This difference in performance between tasks decreased with increasing noise level. Our observations are supported by the corresponding statistics [main effect of task: $F(2, 30) = 4.703$, $p = .017$; $F(2, 30) = 124.211$, $p < .0001$; interaction Task x Noise Level: $F(4, 60) = 29.927$, $p < .0001$]. During the highest noise level, performance was at chance both during the Effector and the Goal Task, as indicated by D' not significantly differing from 0 [Effector Task High Noise Level vs 0: $t(15) = -1.043$, $p = .313$; Goal Task High Noise Level vs 0: $t(15) = 1.821$, $p = .089$]. Moreover, performance did not differ between the Effector and the Goal Task [Effector Task High Noise Level vs Goal Task High Noise Level: $t(15) = -1.588$, $p = .133$] during the highest noise level.

Behavioral Data: Reaction Time. Mean reaction time and standard error (SE) in the Red Dot, Effector, and Goal Task was 464.54 (+/- 10.30), 447.73 (+/- 9.78) and 460.62 (+/- 8.57) ms, respectively [main effect task: $F (2,30) = 3.393$, $p = .047$]. Pairwise comparisons revealed that the Red Dot Task yielded longer response times than the Effector Task [$F(1,15) = 6.465$], whereas none of the remaining comparisons revealed significant differences (all $p > .05$). Mean reaction time was shorter with low (435.05 +/- 9.15 ms) in comparison to the medium (468.46 +/- 10.12 ms) and high (469.38 +/- 8.41 ms) noise levels [main effect noise level: $F(2, 30) = 28.130$, $p < .0001$]. Task and noise level did not interact [$F(4, 60) = 1.725$, $p = .156$].

fMRI Data: Areas involved in action understanding. The areas identified by the contrast [Effector Task, Goal Task] > Red Dot Task (see section ROI definition for details) are shown in Figure 2B. This statistical contrast revealed a region in the orbital part of the left and right inferior frontal gyrus (anterior ventrolateral prefrontal cortex, aVLPFC), the triangular and opercular part of the left and right inferior frontal gyrus (IFG), left dorsal premotor cortex (PMd), bilateral inferior parietal lobe (IPL) and bilateral middle temporal gyrus (MTG). Talairach coordinates of these ROIs can be found in Supplementary Table 1.

fMRI Data: The Effect of Task and Noise Level. Within these ROIs, we extracted the beta estimates of the BOLD response as a function of Task and Noise Level. Since we used the comparison between the Effector and the Goal Task versus the Red Dot Task to identify our regions of interest, we restricted statistical analysis to the comparison between the Effector and the Goal Task (the results of the Red Dot Task in Figures 2C and D are shown for illustrative purposes only).

Figure 2C-E shows the effect of Task and Noise Level as a function of ROI. Overall activation level increases from anterior to posterior ROIs (main effect of ROI: [$F(8, 128) = 35.078$, $p < .0001$]). This observation holds for all three noise levels. The Red Dot Task (red bars) lead to an overall weaker response in comparison to the Effector (gray bars) and the Goal Task (blue bars) (Figure 2C-E). This is an expected result since we selected ROIs accordingly. Importantly, however, in most of the regions, the Goal Task leads to a higher BOLD signal in comparison to the Effector Task [main effect Task: $F(1, 16) = 8.822$, $p = .009$]. In most ROIs, the difference between the Effector and the Goal Task is only observed for the low noise level, i.e. those trials when the stimulus was easy to recognize [(Figure 2C). This observation is supported by the corresponding interaction between Task x Noise Level [$F(2, 32) = 3.911$, $p = .030$]. In IFG and aVLPFC, this difference between the Goal and the Effector Task is present also for the medium noise level (Figure 2D). This difference of the effect of the Noise Level on the effect of the Task between ROIs is supported by the corresponding interaction for the factors ROI x Task x Noise Level [F (9.988,

159.808) = 3.720, pHF < .001]. For the high noise level, none of the regions showed a significant difference between the Effector and the Goal Task (Figure 2E). For an overview of the corresponding statistics, see Table 1.

	df	F	p
ROI	(8, 128)	35.078	< .0001
Task	(1, 16)	8.822	.009
Noise Level	*(2, 32)*	*2.134*	*.135*
ROI x Task	(3.848, 61.575)	14.549	<.0001 (HF)
ROI x Noise Level	(8.794, 140.711)	3.720	<.0001 (HF)
Task x Noise Level	(2, 32)	3.911	.030
ROI x Task x Noise Level	(9.988, 159.808)	3.720	<.0001 (HF)

Table 1*: Results of ANOVA for factors ROI (6) x Task (2) x Noise Level (3). If Mauchly's tests indicated violation of the sphericity assumption, degrees of freedom were adjusted by the Huynh–Feldt procedure (denoted as HF).*

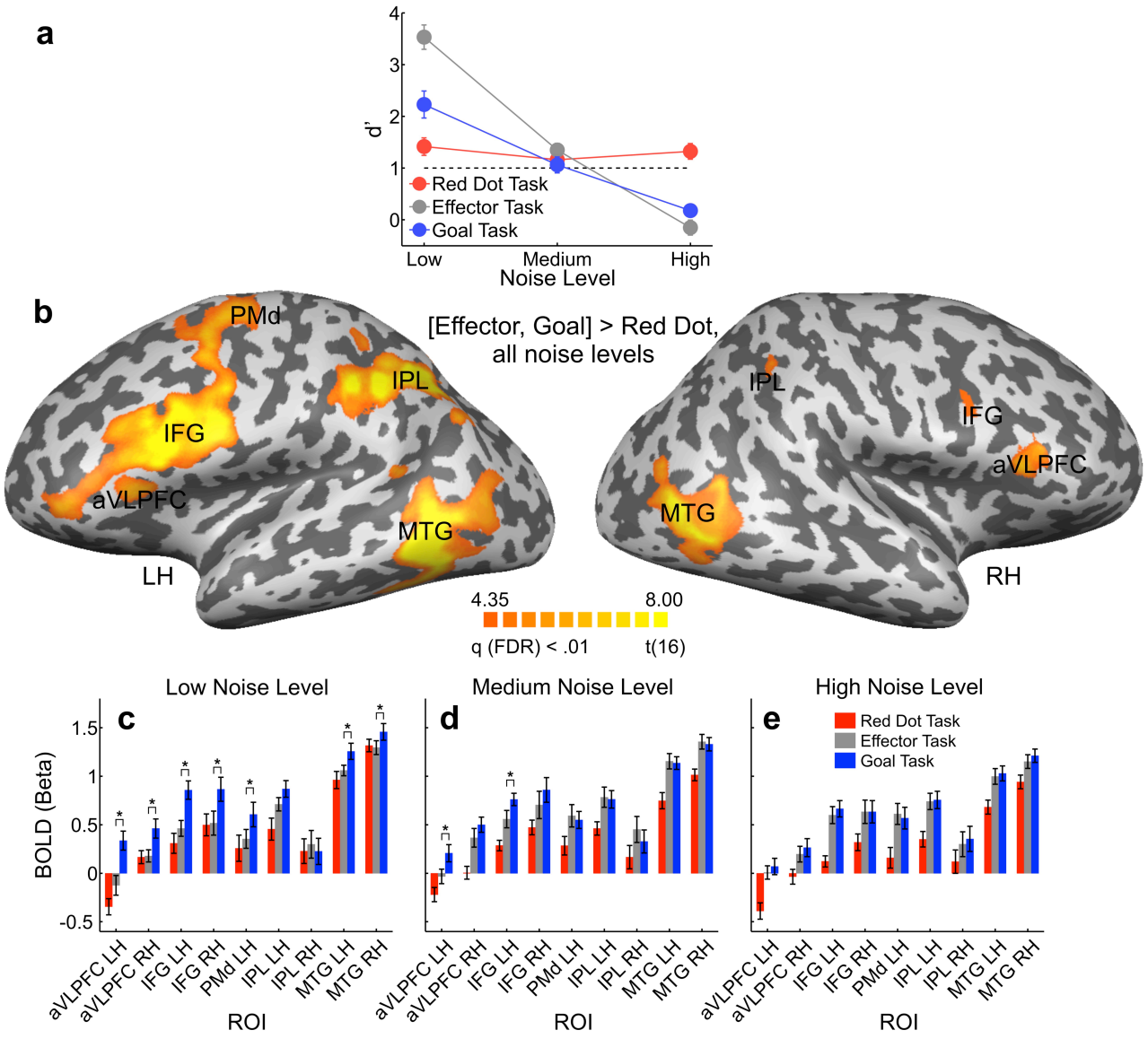


Figure 2. *(a) Behavioral data. Performance, expressed as d0, for N=16 participants (behavioral data from 1 participant were not correctly logged due to technical problems and therefore had to be excluded from the analysis). Error bars show ±SE of the mean. (b) Statistical map revealed by the contrast "[Goal Task, Effector Task]> Red Dot Task" (N=17 participants, random-effects GLM; FDR <0.01; cluster size threshold 4 voxels), projected on top of the segmented and inflated left hemisphere of 1 of the participants. (c) The effect of the task for each ROI for the low noise level. Areas that are sensitive to the difference between the Effector and the Goal Tasks include, but are not restricted to, sensory-motor areas, in particular, the middle temporal gyrus. (d) The effect of the task for the medium noise level. If the task gets harder, the difference between the Goal Task and the Effector Task is restricted to frontal regions (aVLPFC, IFG). (e) The effect of the task for the high noise level. PMd, dorsal premotor cortex; LH, left hemisphere; RH, right hemisphere. Asterisk indicates significant effect (P<0.05) of pairwise comparisons between Effector and Goal Tasks. See text for details.*

Within each ROI, we carried out additional ANOVAs with the factors Task and Noise Level (see Table 2).

	Main Effect of Task		Main Effect of Noise Level		Interaction Task x Noise Level	
	F (1, 16)	p	F (2, 32)	p	F (2, 32)	p
aVLPFC LH	16.878	.001*	.752	.479	5.197	.011*
aVLPFC RH	12.246	.003*	4.214	.024*	2.251	.122
IFG LH	14.756	.001*	.162	.852	5.926	.006*
IFG RH	15.964	.001*	2.256	.121	6.347	.005*
PMd LH	1.508	.237	3.201	.054	4.671	.017*
IPL LH	2.267	.152	.262	.771	2.188	.129
IPL RH	1.090	.312	1.811	.180	1.524	.233
MTG LH	5.206	.037*	6.978	.003*	2.785	.077
MTG RH	9.349	.008*	12.566	<.0001*	2.633	.087

***Table 2**: Statistical details for main effect of task, noise level and interaction noise level x task. Significant values ($p < .05$) are marked by an asterisk.*

This analysis revealed that the effect of task is not restricted to parietal and premotor areas. Both within the left and right middle temporal gyrus and the anterior ventrolateral prefrontal cortex, the Goal Task leads to a higher BOLD signal than the Effector Task, and this effect is modulated by the noise level. Note that the lack of a difference between the Effector and the Goal Task for medium and high noise levels in MTG cannot be due to a ceiling effect: if that were the case, we should expect to see an increase of overall activity with noise level. By contrast, the BOLD response decreases with increasing noise level both in left [main effect noise level: $F(2, 32) = 6.978$, $p = .003$; linear trend: $F(1, 16) = 8.204$, $p = .011$] and right [main effect noise level: $F(2, 32) = 12.566$, $p < .0001$); linear trend: $F(1, 16) = 15.922$, $p = .001$] MTG.

2.8 Discussion

It has been claimed that the action mirroring system enables a direct, "not cognitively mediated" access to action understanding (Fabbri-Destro & Rizzolatti, 2008). In contrast to this view, we show that a task that requires understanding the goal in comparison to the effector of an action does not only recruit parietal and premotor areas involved in action mirroring, but also bilateral MTG and the aVLPFC.

Action mirroring as the outcome of action understanding or vice versa? We observed that the difference between the Goal and the Effector Task in bilateral MTG was restricted to the low noise level condition, i.e. when actions were easy to understand. In contrast, frontal areas also distinguished between the two tasks at an intermediate noise level, i.e. when actions were hard to understand. This suggests an interesting division of labor between temporal and frontal areas, with the MTG potentially being involved in a first pass of the analysis that generates hypotheses about the meaning of an action, while frontal areas, in particular aVLPFC and IFG, might provide additional information when it becomes harder to distinguish between several competing alternatives.

Alternatively, the involvement of bilateral MTG in action understanding might reflect feedback from premotor areas: information enters the IFG through the IPL, and once the action has been understood on the basis of action mirroring, the output is transmitted to bilateral MTG. If IFG was the central hub for action understanding, lesions to IFG, but not MTG should impair action understanding even if the task is easy. In contrast to this prediction, Tranel et al. (2003) reported impaired retrieval of action knowledge not only by lesions in premotor and parietal cortex, but also by lesions in the left posterior MTG. Likewise, using voxel-based lesion symptom mapping (VLSM; Bates et al., 2003), Kalenine et al. (2010) reported that lesions in the MTG, but not in the IFG, were associated with impaired performance in a gesture recognition task. These results are compatible with the idea that MTG provides access to action semantics stored in memory, which

are then fed forward to other areas, including IFG, where this information is matched with the corresponding motor plan, e.g. in preparation of an imitative gesture, and where additional information might be recruited, e.g. on the basis of the knowledge of how these actions are performed. Note that the task used by Kalenine et al. (2010) was relatively easy. This might be the reason why they observed no significant relation between performance and lesions in IFG but in the MTG. Future studies using VLSM or TMS might profit from the parametric variation of task difficulty described in the present study.

The role of MTG and aVLPFC in action understanding. Taken together, our data suggest a specific involvement of areas within the temporal cortex that contain elaborated representations of actions in addition to parietal and premotor areas that have been suggested to be involved in action mirroring. We hypothesize that the MTG might respond to sequences of visual or auditory events and gather evidence in favor of stored semantic representations. This information might then be sent to posterior IFG via the aVLPFC, where this semantic information is then matched with the corresponding motor representation. The suggested interplay between MTG and aVLPFC is supported by a recent study using diffusion-tensor imaging (DTI) and resting state fMRI (A U Turken & Dronkers, 2011) that reported that the left MTG is strongly connected with the orbital part of the IFG, anterior superior temporal gyrus and the superior temporal sulcus. These findings support the idea that the MTG might be seen as a central hub in the network involved in retrieval and storage of action semantics. According to this view, lesions in MTG should result in profound deficits, in line with the findings reported by Kalenine et al. (2010) and Tranel et al. (2003). In line with this view, a meta-analysis on fMRI studies aimed to identify brain regions involved in the storage and retrieval of conceptual knowledge reported that the largest overlap of studies that specifically examined action knowledge was reported in the posterior left MTG and the ventral left supramarginal gyrus in the vicinity of the IPL (Binder, Desai, Graves, & Conant, 2009).

The effect of task difficulty. Our behavioral data show that the Goal Task was more difficult than

the Effector Task at the lowest noise level. Is it possible that the larger BOLD signal in the Goal in comparison to the Effector Task reflects task difficulty rather than differences in the level of action understanding required by the two tasks? We think that this is unlikely: if the areas revealed in our paradigm responded to differences in task difficulty only, the BOLD signal should follow the behavior not only in the Goal and the Effector Task, but also in the Red Dot Task. This is clearly not the case: in the low noise condition, the Red Dot Task leads to worst performance, whereas the Effector Task leads to best performance. The BOLD signal, on the contrary, is highest for the Goal Task, intermediate for the Effector Task and lowest for the Red Dot Task. For the High Noise Level, performance for the Red Dot Task is better in comparison to the Effector and the Goal Task, whereas the Goal and the Effector Task show a higher BOLD signal in comparison to the Red Dot Task also under these circumstances. Taken together, differences in task difficulty between the three tasks cannot explain the BOLD responses we observed.

One may wonder why we observed no difference between the Goal and the Effector Task for the highest noise level. For the most difficult condition, i.e. when 12 out of 13 markers were scrambled, behavioral results (d') showed that participants performed at chance, and performance did not differ between the Effector and the Ball Task. It is therefore not surprising that frontal regions do not distinguish between deep vs shallow action understanding at the highest noise level. In other words, we assume that frontal areas might provide additional information when action understanding becomes harder, but not when it is impossible to understand the action.

Sensitivity to objects. Beauchamp, Lee, Haxby, & Martin (2003) reported that the middle temporal gyrus responds stronger to point-light displays of tool motion in comparison to human motion. Does the difference between the Goal and the Effector Task in the MTG reflect sensitivity to objects, which are relevant in the Goal Task, but not in the Effector Task? Since actions involving a ball were equally likely to occur during the Goal as well as during the Effector Task, a simple explanation on the basis of the presence or absence of an action involving a ball during these two

conditions does not hold. It is possible, however, that participants paid attention to whether or not a ball was involved during the Goal Task, and that this enhanced the signal in object-sensitive neuronal populations, thereby causing a higher response during the Goal Task in comparison to the Effector Task. Such an explanation would predict a higher BOLD signal for actions involving a ball (i.e. throwing a ball, kicking a ball) in comparison to those actions not involving a ball (i.e. punching someone, kicking someone) during the Goal Task, but not during the Effector Task. To test this prediction, we carried out an additional GLM analysis, using the factors Type of Action (actions involving a ball, actions not involving a ball) and Task (Effector, Goal). In order to have enough statistical power for this analysis, we collapsed actions involving the arm and the leg as well as trials with low and medium noise levels, leaving out trials with high noise levels. Within each ROI, we estimated the beta weights for each of these factorial combinations and submitted the resulting values to a repeated-measures ANOVA. The results of this analysis can be seen in Supplementary Figure 1 and Supplementary Table 2. None of the areas followed the predicted pattern. If at all, actions not involving a ball tended to lead to a higher BOLD signal in comparison to actions involving a ball (significant in the right MTG, not significant in all the remaining regions). We are therefore confident that object sensitivity in MTG does not explain the results we observed.

2.9 Conclusions

It has been suggested that parietal and premotor areas that are activated when we perform actions ourselves might be involved in recognizing how an action is performed and what the meaning of that action is. The current study suggests that the meaning of an action is not accessed in the action mirroring network alone, but in additional areas involving the middle temporal gyrus, i.e. an area that binds together information from different modalities and has access to stored memory representations related to these various inputs, possibly through direct interactions with the anterior ventrolateral prefrontal cortex. These results provide important constraints for biologically plausible

models of action understanding and the role of action mirroring.

3. STUDY II: ACTION UNDERSTANDING WITHIN AND OUTSIDE THE MOTOR SYSTEM: A DIFFUSION TENSOR IMAGING STUDY

Abstract

In our previous fMRI study (Lingnau & Petris, 2012) we aimed at identifying cortical areas involved in action understanding. Participants were presented with point-light displays of different human actions and performed tasks that required either understanding the goal (Goal task) or identifying the effector (Effector task) of the observed action. The comparison of these tasks identified several areas, within and most importantly outside the human mirror neuron system such as inferior-frontal gyrus (IFG), inferior-parietal lobule (IPL) and middle temporal gyrus (MTG). Here I aimed to provide evidence about connections between these regions using diffusion tensor imaging (DTI). Particularly I aimed to show which white matter bundles show association between white matter integrity expressed as fractional anisotropy (FA) and variation in performance in tasks used in our previous study. I found associations whithin left superior longitudinal fasciculus (SLF), corticospinal tract (CST) and inferior fronto-occipital fasciculus (IFOF). This finding suggests the existance of two pathways – dorsal and ventral - potentially involved in action understanding.

3.1 Introduction

In our daily life, we constantly have to interpret actions performed by other people. Understanding whether an individual is going to punch or caress us determines how we will react, and being unable to do so can be dangerous under some circumstances. Understanding how actions such as kicking a ball or playing the piano are represented in the brain thus is a fundamental question for cognitive neuroscience.

The topic of action understanding has become very popular thanks to the discovery of mirror neurons - neurons that are active both when an action is observed and executed – in F5 region, which is part of the ventral premotor cortex of the macaque brain (Di Pellegrino et al., 1992). In

humans, this mirror mechanism has been suggested to be the neural mechanism at the core of action understanding, allowing the observer to understand the action through a mechanism of direct – not cognitively mediated – matching between perception and action (Giacomo Rizzolatti & Sinigaglia, 2010).

In humans, the mirror neuron system is thought to be formed by the ventral premotor cortex, including the inferior-frontal gyrus (IFG), and the inferior parietal lobule (IPL) (Rizzolatti & Sinigaglia 2010). In support of the direct matching account, a series of fMRI and TMS studies have been performed demonstrating the involvement of the motor system during action processing. For example, using transcranial magnetic stimulation (TMS), Fadiga et al. (1995) found enhanced motor evoked potentials (MEPs) recorded from different hand muscles of volunteers during the observation of the experimenter grasping an object compared to observation of the same object. Electromyography was also recorded during execution of grasping actions. The authors found enhanced MEPs during action observation compared to the object observation. Moreover, muscles activation induced by TMS during observation was very similar to that involved during execution of the same action. The authors concluded that such a motor facilitation demonstrates the existence of a matching system between action observation and execution in humans similar to the mirror system found in monkeys. Similarly, Buccino et al. (2001), using functional magnetic resonance imaging (fMRI), localized brain areas activated during observation of object and non-object-related actions made by another individual with different effectors (mouth, arm/hand, foot), in the ventral premotor and parietal cortices. Moreover, this activation showed a somatotopic organization. The authors concluded that when an individual is observing another individual performing an action, an internal replica of that action is generated in the premotor cortex and in the case of object-related actions an additional analysis takes place in the parietal lobe. The authors argued that the direct matching system is not only restricted to the ventral premotor cortex, but that it is extended to the parietal lobe as well.

Although the above mentioned studies clearly demonstrate the involvement of the motor system during action observation, they do not tell whether the motor system has a functional role in processing higher order representations of actions such as understanding the goal of an action. As proposed by non-motor theories of action understanding (Mahon & Caramazza, 2008), the activation of the motor system could be the outcome of action understanding rather than the processing of action understanding itself. According to this view, the motor system could receive semantic information from areas outside the hMNS and get activated for action preparation or imitation. In an attempt to find out whether the mirror neuron system processes more abstract representation of actions, many neuroimaging studies have been performed. For example, Hamilton et al. (2006) using fMRI adaptation were interested in finding which areas of the brain were sensitive to the immediate goal of an action such as "reaching to take a cookie or a disk" (with the hand) irrespective of the kinematic of the movements such as the trajectory of the arm used to execute the action. They found the anterior intraparietal sulcus to be sensitive to the immediate goal of an action irrespective of the kinematic of the arm. Similarly Cattaneo et al (2010) found PMv and IPL to be sensitive to pushing or pulling actions irrespective of the effector (hand or foot) and objects (ball or cube) used to perform the actions. By contrast, STS was sensitive to the type of action (pushing and pulling) but only if the action was accompanied with the effector used to perform it. The authors concluded that an abstract processing of observed actions takes place in PMv and IPL while STS shows some abstraction of motor acts but to a lesser degree

In our previous fMRI study (Lingnau & Petris 2012) we also aimed at identifying cortical areas involved in action understanding. Participants were presented with point-light displays of different human actions and performed tasks that required either understanding the goal (Goal task) or identifying the effector (Effector task) of the observed action. We reasoned that both the Effector and Goal tasks would allow action understanding with the Effector task requiring a shallow level of action processing while the Goal task requires a deeper level of action processing. Moreover, as a

control task participants were asked to perform a task where they had to decide if one of the dots composing the point-light display briefly turned red (Red Dot task).

The comparison of these tasks identified several areas, in particular the IFG and IPL, i.e. areas that are typically reported to be involved in the mirror neuron system. However, in contrast to the above mentioned studies, we also identified areas outside the motor system, in particular the middle temporal gyrus (MTG) and the anterior ventrolateral prefrontal cortex (aVLPFC). In line with our results, several studies report MTG as a potential critical node within the neural network involved in processing of conceptual knowledge (M Bedny et al., 2011; Marina Bedny et al., 2008; D Tranel, Damasio, & Damasio, 1997). Our results are problematic for motor theories of action understanding since they assume that the process of action understanding takes place within the motor system. Whereas our results report the involvement of brain regions outside the motor system, the specific role played by these different areas and the connections among them remains unclear. Although an increasing number of studies are focussing on the connectivity – structural and functional – underlying cognitive processes (e.g., Smith, 2012), very little is known about the structural and functional connectivity between brain areas involved in action understanding.

Previous DTI studies reported a connection between MTG and IFG through the left Inferior fronto-Occipital Fasciculus (IFOF), a white matter bundle connecting inferior-lateral and dorso-lateral frontal cortex with posterior temporal cortex and the occipital lobe (Catani, Howard, Pajevic, & Jones, 2002). Moreover, several studies reported IFOF as being part of a semantic system involved in amodal semantic processing (De Zubicaray, Rose, & McMahon, 2011; Duffau et al., 2005), suggesting a ventral pathway responsible for transferring information from amodal semantic areas in the temporal lobe to the mirror neuron system in the frontal lobe (James M Kilner, 2011).

DTI studies rely on the diffusion of water in brain tissue. In a medium of pure liquid with no barriers the diffusion of water molecules is the same in all directions (e.g. it is isotropic). However water molecules diffusivity in the brain is hindered by the microstructure of brain tissue

components such as cell membranes, axons, vessels etc., therefore the direction of water molecules depends by the architecture of the brain microstructure. This dependency is called anisotropy. In brain tissue water molecules diffuse preferentially parallel to axons while they are hindered by barriers such as myelin sheets if diffusion is perpendicular to axons (Chanraud, Zahr, Sullivan, & Pfefferbaum, 2010). Diffusion anisotropy is most commonly quantified via a measure known as fractional anisotropy (FA) (Carlo Pierpaoli & Basser, 1996). Magnetic resonance diffusion tensor imaging is sensitive to anisotropy and therefore it has been developed as a tool for investigating the local properties of brain tissues. Many studies have investigated FA as a marker for white matter integrity in physiological and pathological conditions. Decreased FA values have been found in several pathological conditions such as multiple sclerosis, brain tumors and psychosis (Beaulieu, 2002). Moreover, several studies investigated the correlation between FA values and behavioral performance in patients vs control, in order to investigate the relationship between a structural vs functional damage (for a review of these studies see Chanraud et al., 2010)

Recently this correlation has been investigated in healthy subjects to study how the integrity of specific white matter pathways relate to variation in performance of tasks that should employ those pathways. Such studies aim to better understand the brain network underlying cognitive tasks. For example Johansen-Berg, Della-Maggiore, Behrens, Smith, & Paus (2007), starting from the assumption that inter-individual variation in specific white matter pathways relate to variation in performance of tasks that should employ those pathways, tested whether variation in performance of a bimanual motor task was correlated with the integrity of the white matter in the corpus callosum, measured by FA. The authors found a high correlation between scores obtained in the motor task and FA values in a specific region of the body of the corpus callosum, showing a relationship between behavior and white matter pathways. In another study, De Zubicaray et al. (2011), reported the IFOF as the white matter bundle supporting significant correlation between behavioral performance in semantic tasks and FA values in a population of healthy elderly

participants.

In our previous fMRI study we found activation in posterior MTG and aVLPFC in addition to IFG and IPL. We speculated that the MTG might send semantic information about higher order representations of actions to posterior IFG via the aVLPFC, where this semantic information is then matched with the corresponding motor representation. If the IFOF bundle is involved in semantic processing as suggested by previous studies (Zubicaray et al., 2011; Duffau et al., 2005; Turken et al., 2011), the correlation between behavioral scores in tasks requiring action understanding and FA values should be localized in the IFOF bundle. In the present study we aimed to test this hypothesis.

3.2 Materials and Methods

Participants. 17 healthy subjects (mean age 24.13) participated in this experiment. Two of them were discarded from the analysis either because the behavioral or the DTI data were not acquired properly due to technical problems. The DTI images were acquired at the end of the fMRI experiment reported above (Lingnau & Petris, 2012). All participants were neurologically intact and gave written informed consent for their participation. The experimental procedures were approved by the ethics committee for research involving human subjects at the University of Trento.

Diffusion MRI acquisition. Magnetic resonance images were acquired with a 4 T Bruker Medspec scanner (Bruker Medical, Ettlingen, Germany) using a birdcage-transmit, eight-channel receive head coil (USA Instruments, Inc., Ohio, USA). Each subject underwent a T1-weighted structural image (3D MPRAGE, 1x1x1 mm3, TE = 4 ms; flip angle = 7°, generalized autocalibrating partially parallel acquisition (GRAPPA) factor 2, TI=1020 ms, bandwidth = 150 Hz/pixel, acquisition time = 5 min) optimized for maximal contrast to noise ratio between grey and white matter at 4 T (Papinutto & Jovicich, 2008).

Diffusion weighted images were acquired with a twice refocused 2D SE-EPI sequence (Reese TG, et al., 2003) and the following acquisition parameters: TR=7900 ms, TE = 94 ms, GRAPPA factor

2, voxel size 2 x 2 x 2 mm, b-value 1000 s/mm^2. 5 images without any sensitizing diffusion gradient applied (hereafter called b0) and 30 diffusion weighted images with diffusion gradients applied along unique directions that were defined by an electrostatic repulsion algorithm (Jones DK, 2004; Jones DK et al., 1999) were acquired, with an axial slice acquisition along the x-y plane of the static magnetic field reference frame. A field of view (FOV) of 240 mm2 and 50 contiguous slices enabled to cover the whole brain. A Full-Fourier acquisition was used to reduce cardiac pulsation artifacts (Robson MD et al., 2005).

Behavioral Measurements. Behavioral data were collected inside the scanner during our previous fMRI experiment (Lingnau & Petris, 2012), and DTI acquisition was performed in the same session right after the acquisition of the EPI sequences of the main experiment. Briefly, participants were required to judge point-light displays of 4 different human actions in 3 different tasks. During the Red Dot task participants had to decide whether one of the 13 dots briefly turned red. In the Effector task participants had to decide whether the relevant effector involved in the action was the arm or the leg. During the Goal task participants had to decide whether the action involved the use of a ball or not. The difficulty of the tasks was manipulated by either using the original trajectories (low noise level) or rotating the trajectory of 6 (medium noise level) or 12 (high noise level) out of 13 dots. In all the 3 tasks the response was given by button press. Accuracy was computed using d' statistics (Macmillan & Creelman, 1991) to correct for response bias.

For the DTI analysis we decided to restrict the analysis to the low noise level (the easiest condition) since this condition presented the highest variance. Figure 3 shows the variance associated to each of the 3 task and conditions. As can be seen from the graph, the low noise level condition (blue line) has the highest variance across all three tasks.

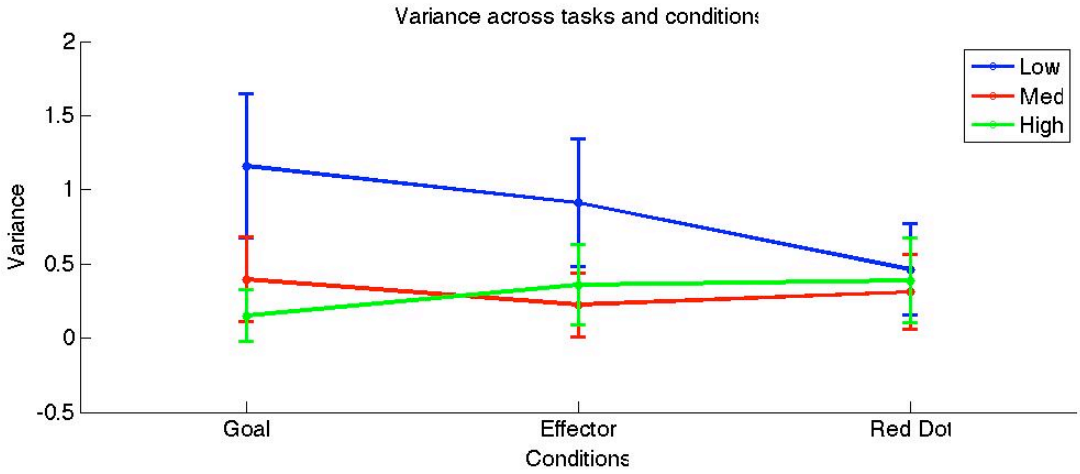


Figure 3. *Variance across the 3 different tasks and conditions employed in Lingnau & Petris (2012). Difficulty of the Red Dot, Effector and Goal tasks were parametrically manipulated by applying different degrees of noise to the point-light displays giving rise to 3 noise level conditions: low (blue line), medium (red line) and high (green line).*

3.3 Preprocessing and Tract-Based Spatial Statistics (TBSS)

Voxelwise statistical analysis of the FA data was carried out using TBSS (Tract-Based Spatial Statistics, (Smith et al., 2006) which is part of the FSL software package (Smith et al., 2004). First, FA images were created by fitting a tensor model to the raw diffusion data using FMRIB's Diffusion Toolbox (FDT), and then brain-extracted using the Brain Extraction Tool (BET) (Smith, 2002). All subjects' FA data were then aligned into a common space using the nonlinear registration tool FNIRT, which uses a b-spline representation of the registration warp field. Next, the mean FA image was created and thinned to create a mean FA skeleton which represents the centres of all tracts common to the group. Each subject's aligned FA data were then projected onto this skeleton and the resulting data were fed into voxelwise cross-subject statistics.

We tested for correlations between behavioral scores, computed as D' Prime, and FA for the Red Dot, Effector and Goal tasks, regressing the behavioral results of each of our three tasks against the whole brain skeletonized FA values and employing a voxel-wise extent threshold of $p = 0.05$ corrected for multiple comparisons using the threshold-free cluster enhancement algorithm (TFCE; Smith & Nichols, 2009) implemented in the "randomise" permutation based inference tool in FSL. However, with a threshold of $p = 0.05$ we did not find any significant results for any of the three

tasks. We therefore decided to report results at a more liberal statistical threshold of 0.15. At this cut-off value we did not find results for the Red Dot task, but we did find a number of clusters within three fibre bundles for the Effector and Goal Task. Supplementary Table 3 (Appendix) reports the number of clusters, the corresponding p value, the MNI coordinates and the associated fiber tracks for both the Effector and Goal tasks.

3.4 Results

Correlation between D' Prime and Effector Task. This regression analysis revealed a positive correlation between D' Prime scores and FA values (Figure 4A) at a threshold of $p = 0.15$ at a whole brain level predominately in the left Superior Longitudinal Fasciculus (SLF), the left Corticospinal Tract (CST) and the Inferior-Frontal Longitudinal Fasciculus (IFOF) as reported by the JHU White-Matter Tractography Atlas implemented in FSL. Figure 5 A, C and E shows the clusters of voxels corresponding to SLF, CST and IFOF, respectively for the Effector Task. No negative correlations were observed.

Correlation between D' Prime and Goal Task. The regression analysis revealed a negative correlation between D' Prime scores and FAs (Figure 4B) at a threshold of $p = 0.15$ predominately in the left SPL and left CST as reported by the JHU White-Matter Tractography Atlas implemented in FSL. Figure 5 B and D shows the clusters of voxels corresponding to the SLF and the CST respectively for the Goal Task. No positive correlations were observed.

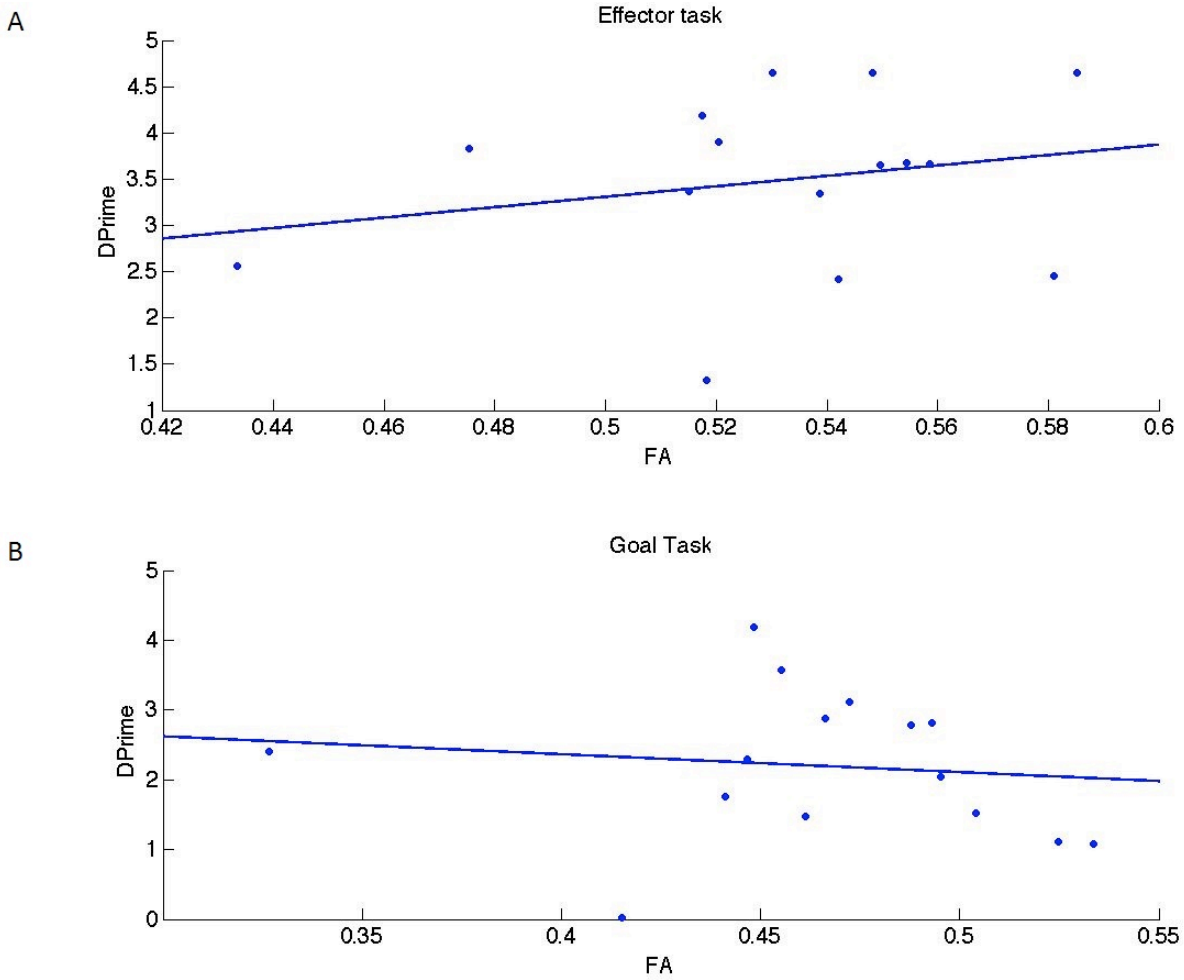


Figure 4. *Correlation between behavioral performance and FA values for 15 subjects.* ***A.*** *Positive correlation between D'prime obtained at the Effector task and FA values.* ***B.*** *Negative correlationbetween D'prime and FA values obtained at the Goal task.*

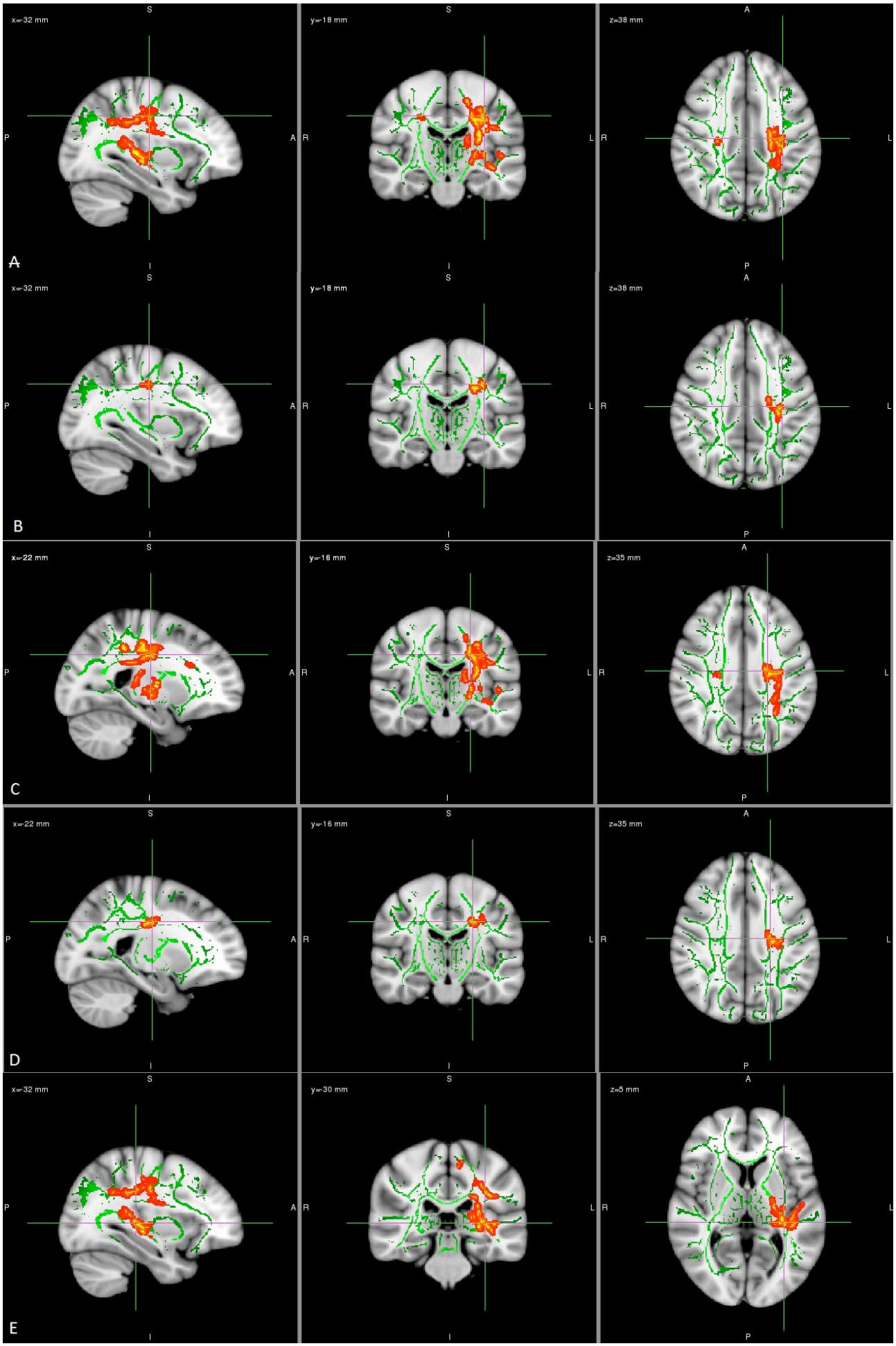

x=-32 mm
y=-18 mm
z=38 mm
A
x=-32 mm
y=-18 mm
z=38 mm
B
x=-22 mm
y=-16 mm
z=35 mm
C
x=-22 mm
y=-16 mm
z=35 mm
D
x=-32 mm
y=-30 mm
z=5 mm
E

Figure 5: *Mean FA skeleton superimposed on MNI brain template for 15 subjects.* ***(A, C, E)****: Effector task. Green cross indicates left SLF* ***(A)****, left CST* ***(C)*** *and left IFOF* ***(E)****.* ***(B, D)****: Goal task. Green cross indicates left SLF* ***(B)*** *and left CST* ***(D)****. From left to right: sagital, coronal and axial views. Enlarged colored clusters represent significant cluster at p = 0.15 corrected for multiple comparisons. SLF: superior longitudinal fasciculus; CST: cortico-spinal tract; IFOF: inferior-frontal occipital fasciculus.*

3.4 Discussion

In the present work we aimed to exploit the correlation between FA values and behavioral performance at tasks requiring action understanding to investigate the white matter bundles involved in action understanding. In our previous study (Lingnau & Petris, 2012), we showed that tasks requiring action understanding recruit areas outside the mirror network, in particular MTG and aVLPFC. It has been suggested that IFOF, a long white matter bundle connecting posterior to anterior parts of the brain, could be part of a semantic network linking MTG to IFG through aVLPFC (Turken & Dronkers, 2011). We therefore decided to test this hypothesis performing a correlation analysis between behavioral scores obtained in our previous study with FA values obtained through DTI. To the best of my knowledge this is the first attempt to investigate the involvement of white matter in action understanding using tract based spatial statistics. It has been shown that fractional anisotropy can be used as a marker of integrity of the white matter since FA values tend to decrease in pathological conditions such as dementia and multiple sclerosis. Furthermore, several studies have shown that this correlation method can be also used in healthy population to investigate the relationship between inter-individual variations in cognitive tasks and white matter integrity.

In the present work I found correlations between behavioral performance at the Effector and Goal task and FA values mainly in the left SLF, CST and IFOF. In the following, I am going to discuss the potential role of each of these fibre bundles for the Effector and the Goal task.

Superior Longitudinal Fasciculus (SLF). The SLF can be characterized as a bundle of white matter tracts including three distinct pathways connecting the temporal to the frontal lobe. The SLF

is best known as arcuate fasciculus (AF) connecting Wernicke's area to Broca's area (Friederici, 2009; Makris et al., 2005). However, in a tractography study, Catani, Jones, & Ffytche (2005) reported that AF is part of a more extensive white matter bundle that forms a pathway from posterior to anterior parts of the brain with several relay stations. This extensive white matter pathway is the superior longitudinal fasciculus. SLF is composed by a long direct segment – the arcuate fasciculus – connecting Wernicke's territory to Broca's territory and an indirect pathway consisting of a posterior segment, connecting Wernicke's territory to parietal areas, and an anterior segment connecting parietal to Broca's territory. As pointed out by Catani et al. (2005) and Catani & Mesulam (2008), Broca's and Wernicke's territories extend well beyond the classical limits of those areas including part of the middle frontal gyrus, inferior precentral and the posterior part of both the superior and middle temporal gyrus, respectively. Although the functional roles of the different segments forming SLF are not completely clear, Catani et al. (2005) suggest a semantic role for the indirect pathway connecting the parietal to frontal areas while the direct pathway would be involved in phonological processing. However, in a meta-analysis of fMRI studies focusing on left hemisphere language areas, Vigneau et al. (2006) did not find support for a semantic role of the parietal cortex. Moreover, recent DTI and fMRI studies (Agosta et al., 2010; Saur et al., 2008) found that sublexical repetition of speech is subserved by the indirect pathway while language comprehension is subserved by a more ventral pathway connecting the middle temporal gyrus to Broca's area. Although our analysis does not tell to which of the different tracts forming the SLF our clusters belong, at a closer eye inspection our clusters seem to be localized below the parietal lobe in a dorsal pathway that could be compatible with the indirect pathway connecting parietal to frontal areas.

How is it possible to reconcile our findings to the connectivity and functional characteristics of the SLF reported so far? First of all it is clear that our experiment requires further investigations to draw more precise conclusions. One hypothesis could be that subjects adopted a subvocalic

repetition strategy to solve the tasks. We cannot rule out that a subvocalic repetition was employed since this strategy is often adopted in behavioral studies. If this was the case, our finding would be compatible with the study of Saur et al. (2008) reporting that the dorsal SLF pathway is involved in such a task. Alternatively, many studies suggest that parietal and frontal areas are involved in coding abstract information of actions such as the goal (Cattaneo et al., 2010; Hamilton & Grafton, 2006). Therefore the finding of SLF in both Effector and Goal task could be in agreement with this hypothesis.

Corticospinal Tract **(CST).** The corticospinal tract is a white matter bundle originating mainly from the primary motor cortex (M1). The neurons of this tract are referred to as pyramidal tract neurons (PTN), because their axons form part of the pyramidal tracts projecting to the spinal cord. M1 is generally considered to initiate and guide movement performance and it is usually not considered to be part of the mirror neuron network mainly because of its executive function (V Gallese, Fadiga, Fogassi, & Rizzolatti, 1996b; M. Iacoboni, 1999). However, a series of TMS and spinal cord exitability studies in humans suggest that M1 and upper motorneurons may be recruited during action observation (Fadiga et al., 1995; Montagna, Cerri, Borroni, & Baldissera, 2005) suggesting that a motor resonance mechanism is at work also in primary motor cortex. Moreover, recently single cell recordings from the pyramidal tract neurons of macaque motor cortex have shown the existence of neurons active during both observation and execution of grasping actions (Kraskov, Dancause, Quallo, Shepherd, & Lemon, 2009; Vigneswaran, Philipp, Lemon, & Kraskov, 2012). Interestingly these studies found that a proportions of these neurons are silent during action observation while discharge during execution. This finding is interesting since it suggests a suppression mechanism that could be recruited during action observation to inhibit the execution of the observed action. We found a correlation between FA values and behavioral performance in the CST for both the Effector and Goal task. This finding is in agreement with the above studies and suggests a motor simulation process of these observed actions.

Infero-fronto Occipital fasciculus (IFOF). The inferior fronto-occipital fasciculus is a long direct connection between occipital and frontal cortex in the human brain (Catani et al., 2002). It has been reported to connect MTG to the orbital part of IFG – the antero-ventrolateral prefrontal cortex (aVLPFC). As noted in the Introduction, previous studies suggested that the IFOF is part of a semantic network for processing conceptual information (De Zubicaray et al., 2011; Duffau et al., 2005). Here we found positive correlation in clusters of voxels localized in the left IFOF bundle for the Effector task only. Although further analyses are needed, this result is in line with the hypotheses of a possible ventral stream responsible for the transmission of semantic information that would reach IFG and match the motor representation of the action. MTG has been already found to be activated during the processing of both motor and non-motor concepts (M Bedny et al., 2011; Bedny et al., 2008) making this region a potential candidate for amodal processing of concepts. Therefore, IFOF could be the pathway involved in transmitting semantic information from MTG to IFG. Problematically, however, we found a correlation between behavioral performance and FA values in IFOF only for the Effector task and not for the Goal task.

3.5 Limitations of the study and Conclusions

Motor theories of action understanding suggest that the parieto-frontal network is involved in processing the goal of an action through a direct mapping of the observed action into the observer's own motor representation of that action. Alternatively, non-motor theories of action understanding claim that the activation of the motor system during action processing tasks is the outcome of a semantic analysis taking place outside the motor system. A possible candidate region for amodal representation of action concepts is the temporal lobe, specifically the MTG. Our results add further information to this picture suggesting which white matter pathways could be involved in the network underlying action understanding. These pathways are the superior longitudinal fasciculus, the inferior-frontal occipital fasciculus and the cortico-spinal tract.

These pathways may have different functions in the neural network of action understanding. CST

receives information from IFG (Catani et al., 2002) where inputs from SLF and IFOF converge. I found the involvement of CST and SLF for both the Effector and Goal task. The task independence of this finding may suggest the activation of a direct and automatic pathway linking IPL-IFG-M1 during action observation that culminates in the simulation of the observed action. Once the motor resonance is active an inhibition mechanisms would refrain the observer from executing the action. This mechanism has been supported by several studies in both humans and monkeys (Fadiga et al., 1995; Baldissera et al., 2001; Montagna et al., 2005; Kraskov et al., 2009; Vigneswaran et al., 2013). In addition, I would like to point out that humans have an automatic tendency to reproduce an observed movement that occurs with or without understanding of the meaning of an action. This tendency is supported by several studies (Barchiesi & Cattaneo, 2012; Brass, Bekkering, Wohlschläger, & Prinz, 2000; Marcel Brass, Bekkering, & Prinz, 2001; Ray & Heyes, 2011). I also found the IFOF, which has been found to be involved in semantic processing. However, problematically, I found IFOF involved in the Effector but not in the Goal task. It is important to note that the experimental design of the behavioral data (Study I) was created to specifically trigger action understanding by the Goal task. Participants were asked to identify the effector involved in the actions compared to identify whether or not a ball was present in the actions the subjects were observing (Goal task). We reasoned that identifying the presence of a ball would have triggered action recognition and action understanding (the "what" of an action) much deeper that the effector task. Therefore it is difficult to explain why we found the IFOF bundle in the Effector rather than the Goal Task. This could be due to a lack of statistical power due to the small number of subjects. Although further analysis are needed to clarify this point, the IFOF bundle could represent a parallel route to converge semantic information from MTG to IFG through aVLPFC (Turken & Dronkers, 2011).

From these preliminary results I speculate that action understanding may be subserve by two routes: a dorsal one passing through SLF and a ventral one passing through IFOF. It is important to note

that both SLF and IFOF converge to IFG, and that IFG sends information to M1. Moreover, Badre & D'Esposito (2009) suggested the presence of a gradient of abstraction of action control in the prefrontal cortex (PFC). Specifically they suggest a rostro-caudal axis in the ventro-lateral prefrontal cortex with more anterior areas such as aVLPFC involved in processing semantic information while more posterior areas such as Brodman area 45 (part of IFG) involved in the decision process about which action has to be perform. Importantly SLF has been found to connect IPL with area 45 of the inferior-frontal gyrus but not to aVLPFC (Petrides, Tomaiuolo, Yeterian, & Pandya, 2012). By contrast IFOF has been suggested to project to aVLPFC and than aVLPFC sends information to IFG. This pattern of connections may suggest an important role for IFG which seems to integrate semantic information in order to match the observed action to the observer's representation of that action. Further evidence is required to sustain the hypotheses stated above. A tractography reconstruction would be helpful for investigating in detail the connections between the different brain regions found in Study I. Moreover, our results do not reach significance, and we found the involvement of the IFOF bundle for the Effector task only. These partial results could be due to different reasons:

a) Lack of variance: although the correlation between structural and functional variations have been performed in healthy population, they are usually performed in patients vs controls, comparing very different variation in behavioral and FA values adding therefore statistical power to the analysis. Even though we chose the condition with the highest variance, it is possible that it was not enough.

b) Crossing fibers: FA values are thought to reflect more efficient information transfer along fibres of similar orientation (e.g., Johansen- Berg et al., 2007). Therefore increasing FA values are thought to represent efficient transfer information. FA, however, is quite sensitive to tissue inhomogeneity from crossing fibers within a voxel (Pierpaoli et al., 2001). Thus, if the white matter voxels are in a region where multiple fiber tracts cross in different directions, FA will be lower, not necessarily

because of reduced fiber integrity but because no single orientation predominates within a voxel (Pierpaoli et al., 2001). It could be that because of crossing fibers, we were less sensitive in some regions to detect effects in the Goal and Effector Task.

4. GENERAL DISCUSSION

In the two studies presented in the current thesis, I aimed to examine which brain areas are modulated by tasks that require understanding the goal of an action compared to identifying the effector of an action (Study I) and which white-matter bundles were involved in these tasks (Study II). In the first study I found that understanding the goal of an action is not restricted to areas within the hMNS as proposed by embodied accounts of action understanding. The results of Study I revealed that tasks that require action understanding do not only recruit parietal and frontal areas but also MTG and aVLPFC. In Study II, I showed that three white matter bundles are involved in action processing: SLF, CST and IFOF. Although the results of the second study are preliminary and the issue about the connections involved in action understanding requires further investigations, taken together these data may suggest some preliminary interpretations.

Regarding Study, I would like to draw some further conclusions. It has been found that MTG is not only involved in semantic processing of actions but it also shows a preference for the grammatical class, verbs versus nouns (e.g. Bedny et al., 2011). One may argue that the preference of MTG for the Goal task with respect to the Effector task is due to a subvocalic repetition of verbs (e.g., "kicking" or "punching" in the Goal Task), in comparison to nouns (e.g., "leg" or "arm" in the Effector task) rather than processing the goal of actions. However this is unlikely since this preference would have been present in the medium and high noise level conditions as well. On the other hand, MTG, as already said above, is also a candidate region for amodal representations of action concepts that will eventually send semantic information to IFG through aVLPFC for the activation of a simulative process that will happen *after* the meaning of the action has been assigned. Our data fit well with this hypothesis in that MTG reveals a higher BOLD response for the Goal task compared to the Effector task. However, this preference is present, along with a

preference for the Goal in comparison to the Effector task shown by IFG, only in the lower but not in the medium or high noise level condition.

It is difficult to explain why MTG does not show a preference for the Goal task also in the medium noise level condition where it is hard but still possible to understand the action. In this latter case one may argue that the preference shown by IFG, but not MTG, reflects the direct matching account for action understanding proposed by mirror neuron proponents, since ventral premotor-IFG has been found repeatedly to be involved in coding the goal of an action. Importantly, however, aVLPFC continues to show a preference for the Goal task also in the medium noise level. aVLPFC has been suggested to be an important centre for the storage of semantic information and it is richly connected to MTG. Taken together, these results show that additional areas such as MTG and aVLPFC are involved in the process of action understanding, providing important constrains to biologically plausible models of action understanding.

Regarding Study II, the aim was to localize white matter bundles involved in action understanding by means of a correlation between behavioral performance obtained in the three tasks used in Study I and FA values, i.e., a measure of the integrity of white matter fibre bundles. The hypothesis was that, given that the IFOF bundle was reported to have a role in brain networks involved in semantic processing, I should have found a significant correlation between performance in the Goal task in comparison to the Effector task in the IFOF bundle. The data, although preliminary, did not confirm this hypothesis. Results suggested an involvement of the SLF and the CST in the Goal task, and the SLF, the CST and the IFOF in the Effector task. Although the Effector task may also be involved in a lower level of action understanding, the critical task in our experiment was the Goal task, i.e., the one that really triggers understanding of the goal of the action.

Apart from methodological limitations (not enough variance in the behavioral performance, small

number of subjects), how can we explain these preliminary findings? CST and SLF could be viewed as direct, automatic pathways connecting parietal (IPL) and frontal (IFG) areas to primary motor cortex. These pathways could give rise to an automatic simulation of the observed action irrespective of the task to be performed. This automatic motor resonance mechanism eventually can be inhibited if the observed action is not going to be produced (Fadiga al., 1995; Baldissera et al., 2001; Montagna et al, 2005; Kraskov et al., 2009; Vigneswaran et al., 2012) An alternative interpretation could be that SLF serves as a subvocalic repetition of the verbs and nouns involved in the tasks (e.g., kicking or punching, arm or leg) adopted by the subjects. Although this hypothesis cannot be ruled out, we think it is unlikely since this option has been already discarded in Study I (see above).

Regarding the IFOF bundle, it links MTG to IFG via aVLPFC, while SLF links IPL to Brodman area 45 (part of IFG) but not aVLPFC. Interestingly, Badre & D'Esposito (2009) suggest a gradient of abstraction in the rostro-caudal axis of the prefrontal cortex with more anterior areas such as aVLPFC being involved in semantic information storage, while more posterior areas such as IFG (Brodman areas 45 and 46) being involved in the decision of the correct action to plan or execute. It seems therefore plausible to assume that IFG may play a role in integrating semantic information in order to decide whether or not the action will be executed. In my view there might be 2 routes underpinning action understanding: one passing through the SLF bundle and the other passing through the IFOF bundle. Both converge to IFG and then IFG sends information to M1 for the simulation of the action that will be inhibited if the action will not be executed. MTG is thought to be an amodal brain region where action concepts are stored irrespective of their sensory and motor properties. In other words, MTG might contain “abstract” representations of actions. A study by Bedny et al. (2011) illustrates this point comparing congenitally blind and sighted individuals in an fMRI study by aurally presenting them verbs and nouns with different degree of visuo-motor

content (i.e., running, thinking, cat, rock). The authors found that MTG was more active for verbs in comparison to nouns in both groups of participants, with equal responses for motor and non-motor content verbs (e.g. running and thinking). The authors concluded that visual experience is not necessary for representing action concepts and that MTG may store amodal representations of concepts. However, recently a very similar study has been published by Ricciardi et al. (2013), where congenitally blind and sighted individuals were aurally presented with the sound of hand actions compared to environmental sounds. Subjects had to recognize the actions while the BOLD response was acquired. The authors found that in both groups the putative hMNS was activated leading the authors to the conclusion that *"the mirror system stores a motor representation of others' actions that can be evoked through supramodal sensory mechanisms"* However the authors also found MTG along with the hMNS. These two studies found similar results in that both found MTG, while the difference relies on the stimulus materials which is represented by words in the study by Bedny et al., (2011) and by action-related sounds in the study by Ricciardi et al. (2013). Taken together, these two studies suggest an important role for MTG since MTG responds to both verbally and aurally-related-action stimuli in both sighted and blind individuals suggesting therefore that MTG may store abstract sensory -independent action concepts. By contrast, in the study by Bedny et al. (2011), no signs of an involvement of the mirror network was found.

4.1 Conclusions

To conclude, what is the role of the sensory-motor system in action knowledge? I think that a certain degree of abstraction is present in the putative hMNS as demonstrated by several studies showing activation of IFG and IPL to goal related actions irrespective of the effector and object used to perform those actions (Hamilton et al., 2006; Cattaneo et al., 2010). However, in my view, action knowledge is not confined to the motor system as demonstrated by Study I and II and

supported by other studies, especially patient studies (Negri et al, 2007; Kalenine et al, 2010). As suggested above, other brain areas such as MTG and aVLPFC may contain a higher level of abstraction not only relative to action concepts, but also in relation to non-sensory-motor related concepts such as thinking, judging etc. that are preferentially activated by language. MTG could represent a central semantic system of all possible concepts abstracted away from the sensory-motor inputs/outputs that instantiate them. According to this view, the semantic system would be linked with the motor system where a lower degree of abstraction is required, as suggested by the studies of Badre & D'Esposito (2009). In line with these studies, an interesting approach in interpreting the literature on action understanding is the Generative Model by Kilner (2011). In this approach, predictions of the goals and intentions of an observed action are generated in a ventral pathway connecting MTG to aVLPFC. Once this information reaches IFG, which supports more concrete representations of actions, the most appropriate goal of the observed action is selected and then it is matched with the observed action in more posterior parts of the PFC.

Further studies are required to clearly understand which of the many suggested models fit the vast amount of data regarding mirror neurons and action understanding. The data reported in Study I and II suggest clearly that other areas outside the hMNS are recruited for tasks requiring action understanding. Moreover, Study II suggests that two parallel routes subserve action understanding, with the ventral one potentially being involved in higher order representation of actions such as generating hypotheses about goals and intention, in agreement with Kilner (2011), while the dorsal pathway might be involved in more concrete representations of action. These two pathways seem to converge in IFG where appropriate responses to the observed action are generated.

5. REFERENCES

Agosta, F., Henry, R. G., Migliaccio, R., Neuhaus, J., Miller, B. L., Dronkers, N. F., Brambati, S. M., et al. (2010). Language networks in semantic dementia. *Brain : a journal of neurology*, *133*(Pt 1), 286–99. doi:10.1093/brain/awp233

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature reviews. Neuroscience*, *10*(9), 659–69. doi:10.1038/nrn2667

Barchiesi, G., & Cattaneo, L. (2012). Early and late motor responses to action observation. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nss049

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*. doi:10.1146/annurev.psych.59.103006.093639

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., & Dronkers, N. F. (2003). Voxel-based lesion-symptom mapping. *Nat Neurosci*, *6*(5), 448–450. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12704393

Beauchamp, M. S., Lee, K. E., Haxby, J. V, & Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, *15*(7), 991–1001. doi:10.1162/089892903770007380

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in biomedicine*, *15*(7-8), 435–55. doi:10.1002/nbm.782

Bedny, M, Caramazza, a, Pascual-Leone, a, & Saxe, R. (2011). Typical Neural Representations of Action Verbs Develop without Vision. *Cerebral Cortex (New York, N.Y. : 1991)*, 1–8. doi:10.1093/cercor/bhr081

Bedny, Marina, Caramazza, A., Grossman, E., Pascual-leone, A., & Saxe, R. (2008). Concepts Are More than Percepts : The Case of Action Verbs. *Brain Research*, *28*(44), 11347–11353. doi:10.1523/JNEUROSCI.3039-08.2008

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*, *19*(12), 2767–2796. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19329570

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*(4), 433–436. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9176952

Brass, M, Bekkering, H., Wohlschläger, a, & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*(2), 124–43. doi:10.1006/brcg.2000.1225

5. References

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*, 3–22.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Buxbaum, L. J., Kyle, K. M., & Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Brain Res Cogn Brain Res*, *25*(1), 226–239. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15996857

Catani, M., Howard, R. J., Pajevic, S., & Jones, D. K. (2002). Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain. *NeuroImage*, *17*(1), 77–94. doi:10.1006/nimg.2002.1136

Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of neurology*, *57*(1), 8–16. doi:10.1002/ana.20319

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex; a journal devoted to the study of the nervous system and behavior*, *44*(8), 953–61. doi:10.1016/j.cortex.2008.04.002

Cattaneo, L., Caruana, F., Jezzini, A., & Rizzolatti, G. (2009). Representation of goal and movements without overt motor behavior in the human motor cortex: a transcranial magnetic stimulation study. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(36), 11134–8. doi:10.1523/JNEUROSCI.2605-09.2009

Cattaneo, L., Sandrini, M., & Schwarzbach, J. (2010). State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. *Cerebral cortex (New York, N.Y. : 1991)*, *20*(9), 2252–8. doi:10.1093/cercor/bhp291

Chanraud, S., Zahr, N., Sullivan, E. V, & Pfefferbaum, A. (2010). MR diffusion tensor imaging: a window into white matter integrity of the working brain. *Neuropsychology review*, *20*(2), 209–25. doi:10.1007/s11065-010-9129-7

Chong, T. T.-J., Cunnington, R., Williams, M. a, Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology : CB*, *18*(20), 1576–80. doi:10.1016/j.cub.2008.08.068

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition. Attention & Performance XXII.* (pp. 435–459). Oxford: Oxford University Press.

De Zubicaray, G. I., Rose, S. E., & McMahon, K. L. (2011). The structure and connectivity of semantic memory in the healthy older adult brain. *NeuroImage*, *54*(2), 1488–94. doi:10.1016/j.neuroimage.2010.08.058

5. References

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study, *91*, 176–180.

Dinstein, I., Hasson, U., Rubin, N., & Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of neurophysiology*, *98*(3), 1415–27. doi:10.1152/jn.00238.2007

Duffau, H., Gatignol, P., Mandonnet, E., Peruzzi, P., Tzourio-Mazoyer, N., & Capelle, L. (2005). New insights into the anatomo-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. *Brain : a journal of neurology*, *128*(Pt 4), 797–810. doi:10.1093/brain/awh423

Fabbri-Destro, M., & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology (Bethesda, Md.)*, *23*, 171–9. doi:10.1152/physiol.00004.2008

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, *73*(6), 2608–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7666169

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science (New York, N.Y.)*, *308*(5722), 662–7. doi:10.1126/science.1106138

Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends in cognitive sciences*, *13*(4), 175–81. doi:10.1016/j.tics.2009.01.001

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: characterizing differential responses. *Neuroimage*, *7*(1), 30–40. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9500830

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological cybernetics*, *104*(1-2), 137–60. doi:10.1007/s00422-011-0424-z

Gallese, V, Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain : a journal of neurology*, *119 (Pt 2*, 593–609. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8800951

Gallese, Vittorio, & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, *22*(3), 455–79. doi:10.1080/02643290442000310

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *NeuroImage*, *35*(4), 1674–84. doi:10.1016/j.neuroimage.2007.02.003

Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*(4), 1133–7. doi:10.1523/JNEUROSCI.4551-05.2006

5. References

Henson, R. (2006). Forward inference using functional neuroimaging: dissociations versus associations. *Trends in cognitive sciences*, *10*(2), 64–9. doi:10.1016/j.tics.2005.12.005

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of cognitive neuroscience*, *21*(7), 1229–43. doi:10.1162/jocn.2009.21189

Hickok, G., & Hauser, M. (2010). (Mis)understanding mirror neurons. *Current biology : CB*, *20*(14), R593–4. doi:10.1016/j.cub.2010.05.047

Iacoboni, M. (1999). Cortical Mechanisms of Human Imitation. *Science*, *286*(5449), 2526–2528. doi:10.1126/science.286.5449.2526

Iacoboni, M. (2002). Understanding Others: Imitation, Language, Empathy. Retrieved from http://www.sscnet.ucla.edu/CBD/bios/royaumont.pdf

Johansen-Berg, H., Della-Maggiore, V., Behrens, T. E. J., Smith, S. M., & Paus, T. (2007). Integrity of white matter in the corpus callosum correlates with bimanual co-ordination skills. *NeuroImage*, *36 Suppl 2*, T16–21. doi:10.1016/j.neuroimage.2007.03.041

Kalenine, S., Buxbaum, L. J., & Coslett, H. B. (2010). Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain*, *133*(11), 3269–3280. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20805101

Kilner, J M, Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *J Neurosci*, *29*(32), 10153–10159. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19675249

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in cognitive sciences*, *978*, 1–6. doi:10.1016/j.tics.2011.06.005

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron*, *64*(6), 922–30. doi:10.1016/j.neuron.2009.12.010

Lestou, V., Pollick, F. E., & Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *Journal of cognitive neuroscience*, *20*(2), 324–41. doi:10.1162/jocn.2008.20021

Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence. *Proceedings of the National Academy of Sciences of the United States of America, 106*(24).

Lingnau, A., & Petris, S. (2012). Action Understanding Within and Outside the Motor System: The Role of Task Difficulty. *Cerebral Cortex*. doi:10.1093/cercor/bhs112

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge, U.K.: Cambridge University Press.

5. References

Mahon, B. Z., & Caramazza, A. (2005). The orchestration of the sensory-motor systems: Clues from Neuropsychology. *Cognitive neuropsychology*, *22*(3), 480–94. doi:10.1080/02643290442000446

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology, Paris*, *102*(1-3), 59–70. doi:10.1016/j.jphysparis.2008.03.004

Makris, N., Kennedy, D. N., McInerney, S., Sorensen, a G., Wang, R., Caviness, V. S., & Pandya, D. N. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cerebral cortex (New York, N.Y. : 1991)*, *15*(6), 854–69. doi:10.1093/cercor/bhh186

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby; James V. (1996). Neural Correlates of Category-Specific Knowledge. *Nature*, *379*.

Montagna, M., Cerri, G., Borroni, P., & Baldissera, F. (2005). Excitability changes in human corticospinal projections to muscles moving hand and fingers while viewing a reaching and grasping action. *The European journal of neuroscience*, *22*(6), 1513–20. doi:10.1111/j.1460-9568.2005.04336.x

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology : CB*, *20*(8), 750–6. doi:10.1016/j.cub.2010.02.045

Negri, G. a L., Rumiati, R. I., Zadini, A., Ukmar, M., Mahon, B. Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. *Cognitive neuropsychology*, *24*(8), 795–816. doi:10.1080/02643290701707412

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Viewpoint (in)dependence of action representations: an MVPA study. *Journal of cognitive neuroscience*, *24*(4), 975–89. doi:10.1162/jocn_a_00195

Papeo, L., Negri, G. a L., Zadini, A., & Rumiati, R. I. (2010). Action performance and action-word understanding: evidence of double dissociations in left-damaged patients. *Cognitive neuropsychology*, *27*(5), 428–61. doi:10.1080/02643294.2011.570326

Pazzaglia, M., Smania, N., Corato, E., & Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *J Neurosci*, *28*(12), 3030–3041. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18354006

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J Exp Biol*, *146*, 87–113. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2689570

Perrett, D. I., Mistlin, A. J., Harris, M. H., & Chitty, A. J. (1990). Understanding the Visual Appearance and Consequence of Hand Actions. In M. A. Goodale (Ed.), *Vision and action: the control of grasping* (pp. 163–180). Norwood, NJ: Ablex Publishing.

Petrides, M., Tomaiuolo, F., Yeterian, E. H., & Pandya, D. N. (2012). The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. *Cortex; a journal devoted to the study of the nervous system and behavior*, *48*(1), 46–57. doi:10.1016/j.cortex.2011.07.002

Pierpaoli, C., Barnett, A., Pajevic, S., Chen, R., Penix, L. R., Virta, a, & Basser, P. (2001). Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *NeuroImage*, *13*(6 Pt 1), 1174–85. doi:10.1006/nimg.2001.0765

Pierpaoli, C., & Basser, P. J. (1996). Toward a Quantitative Assessment of Diffusion Anisotropy.

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental science*, *14*(1), 92–105. doi:10.1111/j.1467-7687.2010.00961.x

Ricciardi, E., Handjaras, G., Bonino, D., Vecchi, T., Fadiga, L., & Pietrini, P. (2013). Beyond motor scheme: a supramodal distributed representation in the action-observation network. *PloS one*, *8*(3), e58632. doi:10.1371/journal.pone.0058632

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain research. Cognitive brain research*, *3*(2), 131–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8713554

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual review of neuroscience*, *27*, 169–92. doi:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews. Neuroscience*, *11*(4), 264–74. doi:10.1038/nrn2805

Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature reviews. Neuroscience*, *5*(10), 813–9. doi:10.1038/nrn1521

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., et al. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 18035–40. doi:10.1073/pnas.0805234105

Saygin, A. P., Wilson, S. M., Hagler Jr., D. J., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *J Neurosci*, *24*(27), 6181–6188. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15240810

Schwarzbach, J. (2011). A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behav Res Methods*. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&l

ist_uids=21614662

Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, *17*(3), 143–55. doi:10.1002/hbm.10062

Smith, S. M. (2012). The future of FMRI connectivity. *NeuroImage*, *62*(2), 1257–66. doi:10.1016/j.neuroimage.2012.01.022

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*(4), 1487–1505. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16624579

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208–19. doi:10.1016/j.neuroimage.2004.07.051

Smith, S. M., & Nichols, T. E. (2009). Threshold-Free Cluster Enhancement : Addressing problems of smoothing , threshold dependence and localisation in cluster inference FMRIB Technical Report TR08SS1, *1*, 1–20.

Tranel, D, Damasio, H., & Damasio, a R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*, *35*(10), 1319–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9347478

Tranel, Daniel, Kemmerer, D., Adolphs, R., Damasio, H., & Damasio, A. (2003). Neural Correlates of Conceptual Knowledge For Actions. *Cognitive neuropsychology*, *20*, 409–432.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in systems neuroscience*, *5*(February), 1. doi:10.3389/fnsys.2011.00001

Umiltà, M. A, Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, a, et al. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(6), 2209–13. doi:10.1073/pnas.0705985105

Umiltà, M. A, Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing. a neurophysiological study. *Neuron*, *31*(1), 155–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11498058

Urgesi, C., Candidi, M., Ionta, S., & Aglioti, S. M. (2007). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nat Neurosci*, *10*(1), 30–31. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17159990

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the

human motor system. *J Neurosci*, *26*(30), 7942–7949. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16870739

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–32. doi:10.1016/j.neuroimage.2005.11.002

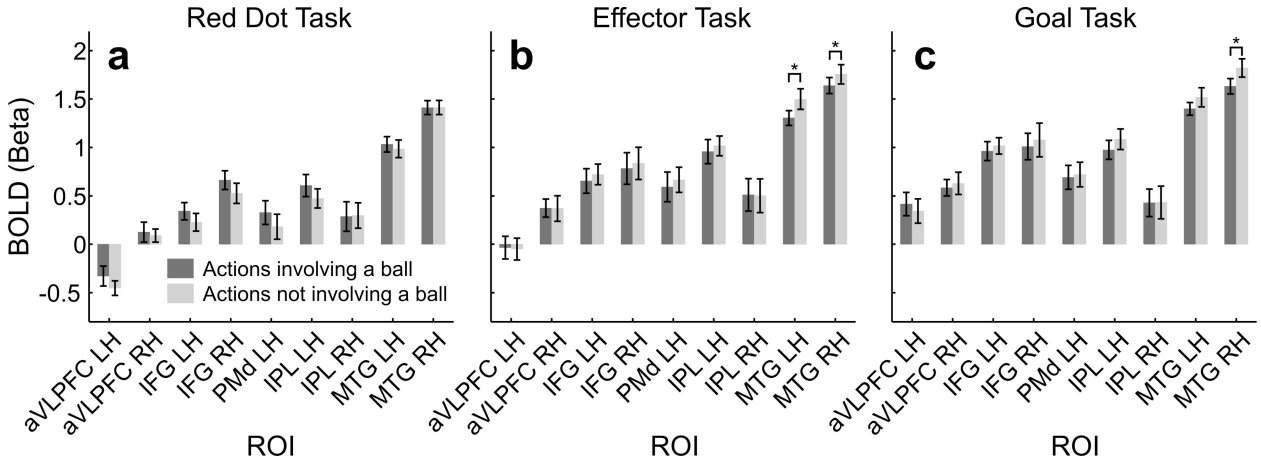
Vigneswaran, G., Philipp, R., Lemon, R. N., & Kraskov, A. (2012). M1 Corticospinal Mirror Neurons and Their Role in Movement Suppression during Action Observation. *Current biology : CB*, 1–8. doi:10.1016/j.cub.2012.12.006

Zaitsev, M., Hennig, J., & Speck, O. (2004). Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Magn Reson Med*, *52*(5), 1156–1166. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15508146

Zeng, H., & Constable, R. T. (2002). Image distortion correction in EPI: comparison of field mapping with point spread function mapping. *Magn Reson Med*, *48*(1), 137–146. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12111941

6. APPENDIX

6.1 Appendix Study I



Supplementary Figure 1*: The BOLD signal as a function of task (A: Red Dot Task, B: Effector Task, C: Goal Task) and action type (involving a ball versus not involving a ball, collapsing over actions performed with the arm and the leg), separately for each ROI identified in the main analysis. Asterisks indicates significant effects of pairwise comparisons between actions involving a ball and actions not involving a ball within those ROIs that showed a significant main effect of action type ($p < .05$) or a significant interaction between action type and task (see Supplementary Table 2). Actions involving a ball do not lead to a higher BOLD signal than actions not involving a ball. The only area revealing a main effect of action type was bilateral MTG during the Effector Task and the right MTG during the Goal Task, where actions not involving a ball led to a higher BOLD signal in comparison to actions involving a ball.*

	Talairach coordinates			Volume
	x	y	z	mm^3, n voxels
aVLPFC LH	-42	35	10	15.2 (3500)
aVLPFC RH	48	28	18	10.7 (1217)
IFG LH	-42	9	25	22.9 (10480)
IFG RH	43	9	27	7.9 (500)
PMd LH	-27	-3	52	17.6 (5476)
IPL LH	-39	-43	37	21.7 (10193)
IPL RH	39	-34	39	7.5 (428)
MTG LH	-45	-58	-2	21.5 (9898)
MTG RH	43	-58	2	19.2 (7028)

Supplementary Table 1*: Talairach coordinates and size of regions of interest revealed by the contrast [Goal Task All Noise Levels, Effector Task All Noise Levels] > Red Dot Task All Noise Levels.*

	Main Effect of Task			Main Effect of Action Type		Interaction Task x Action Type		
	F (2, 32)		p	F (1, 16)	p	F (2, 32)		p
aVLPFC LH	26.050	*	<.0001	1.305	.270	.450		.642
aVLPFC RH	21.371	*	<.0001	.001	.974	.176		.840
IFG LH	28.462	*	<.0001	.002	.967	1.349		.274
IFG RH	10.252	*	<.0001	.007	.932	1.517		.235
PMd LH	16.305	*	<.0001	.054	.819	1.913		.164
IPL LH	38.509	*	<.0001	.045	.835	2.094		.140
IPL RH	4.248	*	.023	.001	.974	.011		.989
MTG LH	34.143	*	<.0001	2.300	.149	3.544	*	.041

MTG RH	23.700	*	<.0001	5.760	*	.029	3.532	* .041

Supplementary Table 2*: Statistical details for main effect of task, action type and interaction task x action type. Significant values (p < .05) are marked by an asterisk.*

6.2 Appendix Study II

Clusters showing relationships between FA and behavioral performance for the Effector and Goal Tasks

Effector Task				Goal Task			
Cluster Index	p-value corrected using TFCE 1-p (.85 -.92)	MNI (x, y, z)	JHU White-Matter atlas	Cluster Index	p-value corrected using TFCE 1-p (.85 -.86)	MNI (x, y, z)	JHU White-Matter atlas
6	..91	-28 -22 37	SLF LH	1	.86	-28-22 37	SLF LH
6	..92	-29 -19 37	SLF LH	1	.86	-23-22 33	CST LH
6	..92	-25 -20 35	CST LH	1	.86	-19 37 24	SPL LH
6	.92	-23 -21 33	CST LH	1	.86	-24-19 32	CST LH
6	.92	-24 -19 32	CST LH	1	.86	-28-25 38	SPL LH
6	.92	-29 -26 40	SLF LH	1	.86	-25-20 37	CST LH
6	.91	-28 -25 38	SLF LH	1	.86	-25-20 35	CST LH
6	.91	-24 -24 44	CST LH	1	.86	-21-16 34	CST LH
6	.91	-28 -23 43	SLF LH	1	.85	-26-18 34	CST LH
6	.91	-26 -24 43	CST LH	1	.85	-28-28 39	SPL LH
6	.903	-31 -30 68	IFOF LH				
5	.86	-12 21 20	ATL LH				
5	.86	-11 21 17	ATL LH				
5	.86	-81 8 17	ATL LH				
5	.86	-9 22 19	ATL LH				
5	.86	-14 27 15	ATL LH				

6. Appendix

		Effector Task		Goal Task
5	.86	-3 17 14	ATL LH	
4	.86	27 -23 39	IFOF RH	
4	.85	26 -19 36	CST RH	
3	.85	23 -20 34	CST RH	
2	.85	-14 10 28	CST RH	
1	.85	-6 -15 13	ATL LH	

Supplementary Table 3. *MNI Coordinates are shown for Effector and Goal Tasks. SLF Superior Longitudinal Fasciculus; CSP Corticospinal Tract; ATL Anterior thalamic radiation; IFOF Iferior fronto-occipital fasciculus; LH Left Hemisphere; RH Right Hemisphere.*